Health Technology Assessment 2012; Vol. 16: No. 4
ISSN 1366-5278

# Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation

M Stevenson, M Lloyd-Jones, MY Morgan and R Wong

February 2012
10.3310/hta16040

**Health Technology Assessment**
**NIHR HTA programme**
**www.hta.ac.uk**



**How to obtain copies of this and other HTA programme reports**

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

– fax (with **credit card details**)
– post (with **credit card details** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

**Contact details are as follows:**

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for 'HTA Payment Services' (out-of-hours answer-phone service)

Fax: 0845 812 4001 – put 'HTA Order' on the fax header

**Payment methods**

*Paying by cheque*
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

*Paying by credit card*
You can order using your credit card by phone, fax or post.

**Subscriptions**

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

**How do I get a copy of *HTA on DVD*?**

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

---

The website also provides information about the HTA programme and lists the membership of the various committees.

# Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation

M Stevenson,[1]* M Lloyd-Jones,[1] MY Morgan[2] and R Wong[1]

[1]The University of Sheffield, School of Health and Related Research (ScHARR), Sheffield, UK
[2]The Centre for Hepatology, University College London Medical School, London, UK

*Corresponding author

**Declared competing interests of the authors**: none

Published February 2012
DOI: 10.3310/hta16040

This report should be referenced as follows:

Stevenson M, Lloyd-Jones M, Morgan MY, Wong R. Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation. *Health Technol Assess* 2012;**16**(4).

*Health Technology Assessment* is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (*SciSearch*®) and *Current Contents*®/Clinical Medicine.

# NIHR Health Technology Assessment programme

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

**Criteria for inclusion in the HTA journal series**

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned by the HTA programme as project number 09/62/01. The contractual start date was in December 2009. The draft report began editorial review in July 2010 and was accepted for publication in March 2011. As the funder, by devising a commissioning brief, the HTA programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.



ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

ISSN 2046-4932 (DVD)



Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.
Printed on acid-free paper in the UK by the Charlesworth Group.

# Abstract

## Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease: a systematic review and economic evaluation

M Stevenson,[1]* M Lloyd-Jones,[1] MY Morgan[2] and R Wong[1]

[1]The University of Sheffield, School of Health and Related Research (ScHARR), Sheffield, UK
[2]The Centre for Hepatology, University College London Medical School, London, UK

*Corresponding author

**Background:** Excessive alcohol consumption may lead to the development of alcohol-related liver disease (ALD). Liver biopsy may be used in patients with suspected ALD to confirm the diagnosis, exclude other or additional liver pathologies, and provide accurate staging of the degree of liver injury in order to enable the prediction of prognosis and inform treatment decisions. However, as it is an invasive procedure that carries the risk of morbidity and mortality, current UK guidance recommends that biopsy is not required to confirm the diagnosis in patients with a high clinical suspicion of ALD in whom blood tests have excluded other causes of liver disease, unless it is necessary to confirm a diagnosis of acute alcoholic hepatitis in order to inform specific treatment decisions.

**Objectives:** To evaluate the diagnostic accuracy, cost-effectiveness, and effect on patient outcomes of four non-invasive tests for liver fibrosis [the Enhanced Liver Fibrosis (ELF™) test (Siemens Healthcare Diagnostic Inc., Tarrytown, NY, USA), FibroTest (BioPredictive, Paris, France), FibroMAX (BioPredictive, Paris, France) and transient elastography (FibroScan®; produced by EchoSens, Paris, France and distributed in the UK by Artemis Medical Ltd, Kent, UK)] in patients suspected of having ALD.

**Data sources:** A systematic review was undertaken to identify studies reporting the diagnostic and prognostic accuracy of the ELF test, FibroTest, FibroMAX, and FibroScan for the identification of liver fibrosis and associated conditions in patients with suspected ALD. The following databases were searched in January 2010: MEDLINE (from 1950 to January 2010), MEDLINE In-Process & Other Non-Indexed Citations (from 1950 to January 2010), EMBASE (from 1980 to January 2010), Cochrane Database of Systematic Reviews (from 1996 to January 2010), Cochrane Central Register of Controlled Trials (from 1898 to January 2010), Cochrane Methodology Register (from 1904 to January 2010), Database of Abstracts of Reviews of Effects (from 1995 to January 2010), HTA Database (from 1995 to January 2010), NHS Economic Evaluation Database (from 1995 to January 2010), Cumulative Index to Nursing and Allied Health Literature (from 1982 to January 2010), Web of Knowledge and Science Citation Index (from 1969 to January 2010).

**Review methods:** Study quality was assessed using the QUADAS (Quality Assessment of Diagnostic Accuracy Studies) checklist. Owing to the heterogeneity of the studies, no formal meta-analysis was undertaken. A de novo mathematical model was constructed to estimate the incremental costs and incremental quality-adjusted life-years (QALYs) associated with alternative strategies compared with a biopsy-all strategy. The tests are assessed first as a replacement for liver biopsy, and secondly as an additional test prior to liver biopsy. Thirty-six scenarios were assessed for each non-invasive test strategy, which

varied the sensitivity of biopsy, the anxiety associated with biopsy, sensitivity and specificity values and whether or not the biopsy was percutaneous or transjugular. For each scenario, threshold levels were reported where biopsying all patients was more cost-effective than the strategy for two parameters (the decreased level of abstinence associated with the strategy compared with biopsying all and the level of incidental QALY gain associated with biopsy).

**Results:** No studies were identified that specifically assessed the ELF test, although a study was identified that evaluated the diagnostic accuracy of the European Liver Fibrosis Test (essentially, the ELF test with the addition of age to the algorithm) compared with biopsy. Three studies of FibroTest, no relevant studies of FibroMax, and six studies of FibroScan assessing accuracy compared with biopsy in patients with known or suspected alcohol-related liver disease were identified. In all studies, the number of patients with suspected ALD was small, meaning that the estimated sensitivities and specificities were not robust. No conclusive estimate of the cost per QALY of each non-invasive test could be provided. Scenarios exist in which each of the strategies analysed is more cost-effective than biopsying all patients and, in contrast, scenarios exist in which each strategy is less cost-effective than biopsying all patients.

**Limitations:** Study selection and data analysis were undertaken by one reviewer.

**Conclusions:** No conclusive result can be provided on the most cost-effective strategy until further data are available. A large number of parameters require data; however, the following are selected as being of most importance: (1) the sensitivity and specificity of each non-invasive liver test (NILT) against biopsy at validated and pre-selected cut-off thresholds; (2) the influence of potential confounding variables such as current drinking behaviour and the degree of hepatic inflammation on the performance of NILTs; and (3) the likelihood, and magnitude, of decreases in abstinence rates associated with a diagnosis of significant ALD by diagnostic modality and the incidental gains in QALYs that may be associated with biopsy.

**Funding:** The National Institute for Health Research Technology Assessment programme.

# Contents


**Glossary** vii

**List of abbreviations** xi

**Executive summary** xiii

**1. Background** 1
Description of health problem 1
Current service provision 10
Description of technology under assessment 11

**2. Definition of the decision problem** 19
Decision problem 19
Overall aims and objectives of assessment 22

**3. Cost-effectiveness: model structure and methodology** 23
The strategies analysed 24

**4. Assessment of clinical effectiveness** 29
Methods for reviewing effectiveness 29
Results 31

**5. Cost-effectiveness: model parameters** 63
Discount rates 63
The sensitivity and specificity of non-invasive liver tests used within the economic model 63
The prevalence of cirrhosis in the defined population 66
The costs of biopsy and each non-invasive liver test 66
Adverse events related to each diagnostic test 67
The proportion of tests that will produce results that cannot be used 67
The outcomes associated with each final node within the economic model for diagnosis of cirrhosis 68
The proportion of patients without cirrhosis who continue to drink heavily that will develop cirrhosis 70
The benefit of biopsy in identifying liver disease that is not alcohol-related liver disease-related and confirming alcohol-related liver disease 72

**6. Cost-effectiveness model: results** 73
Threshold analysis regarding the rates of abstinence 75
Threshold analysis regarding the potential benefits associated with a biopsy 80
Conclusions drawn from the cost-effectiveness results 84

**7. Discussion** 87

**Acknowledgements** 89

**References** 91

**Appendix 1** Categorisation of disease progression as identified by liver biopsy **105**

**Appendix 2** The Alcohol Use Disorders Identification Test **107**

**Appendix 3** Liver biopsy: systematic review of adverse events **109**

**Appendix 4** Systematic review of the adverse effects of liver biopsy: search strategies **119**

**Appendix 5** Assessment of clinical effectiveness and cost-effectiveness, adverse effects, and quality of life: search strategies **121**

**Appendix 6** QUADAS: details of criteria for scoring studies **131**

**Appendix 7** Excluded studies **133**

**Appendix 8** Diagnostic venepuncture: systematic review of adverse events **135**

**Appendix 9** Systematic review of the adverse effects of venepuncture: search strategies **143**

**Appendix 10** The results from the cost-effectiveness analyses **145**

**Appendix 11** Project protocol **153**

**Health Technology Assessment programme** **169**

# Glossary

**Alcoholic fatty liver (hepatic steatosis)** Accumulation of excess fat in the liver cells as a result of excess alcohol consumption.

**Alcoholic hepatitis** An inflammatory condition involving liver damage.

**Ascites** Excessive accumulation of fluid in the peritoneal cavity.

**Banding** Use of elastic bands to tie off blood vessels.

**Beta-blocker** A drug used to lower blood pressure. Propranolol, one of the beta-blockers, is used to treat portal hypertension and so to help prevent bleeding from oesophageal varices in patients with cirrhosis.

**Body mass index** A measure of weight for height obtained by dividing a person's weight in kilograms by the square of their height in metres.

**Cirrhosis** Advanced fibrosis.

**Cohort study** A study that follows groups of people with and without the condition of interest over time to study outcomes.

**Compensated cirrhosis** Cirrhosis with no associated complications.

**Corticosteroids** Drugs that reduce inflammation and can be used to treat severe alcoholic hepatitis.

**Cross-sectional study** A study that measures the determinants of health in a population at a point in time or over a short period of time; in the context of diagnostic studies, a study that examines the accuracy of a test by comparison with a reference standard, rather than by following people who have undergone the test over time to compare their health outcomes with the test results.

**Decompensated cirrhosis** Cirrhosis associated with complications such as ascites, peripheral oedema, jaundice, variceal bleeding, and hepatic encephalopathy.

**Ethanol** Pure alcohol.

**Fibrosis** Fibrous scar tissue.

**Haematocrit** The percentage of red blood cells in a blood sample.

**Harmful drinking** An established pattern of drinking at a level where damage to physical or mental health is considered likely.

**Hazardous drinking** An established pattern of drinking that carries a risk of physical or psychological harm.

**Hepatic encephalopathy** Symptom complex of neuropsychiatric abnormalities observed in patients with acute and chronic liver disease, and mainly affecting mental and motor function.

**Hepatic portal vein** The vein that drains blood from the gastrointestinal tract and spleen to the liver.

**Hepatic steatosis** Accumulation of excess fat within the liver cells.

**Hepatic venous pressure gradient** The pressure gradient within the portal system.

**Hepatocellular carcinoma** Primary cancer of the liver, generally as a consequence of either viral hepatitis or excessive alcohol consumption.

**High-Risk Alcoholism Relapse scale** A scale designed to estimate the risk of alcoholism relapse following evaluation.

**Histological** Related to the microscopic structure of tissue.

**Jaundice** Yellow discoloration of the skin and eyes.

**Kilopascal (kPa)** Metric unit of pressure.

**Laparoscopic** Performed through small incisions in the abdomen.

**Liver biopsy** Removal of a small sample of liver tissue using a hollow needle for examination in the laboratory.

**Liver fibrosis** Formation of excessive fibrous scar tissue in the liver.

**Maddrey score** A measure of severity of alcoholic hepatitis.

**Metabolic syndrome** A combination of central obesity with any two of raised triglycerides, reduced high-density lipoprotein (HDL) cholesterol, raised blood pressure or raised fasting plasma glucose/previously diagnosed type 2 diabetes.

**Non-alcoholic fatty liver disease** Liver disease characterised by the accumulation of fat in the liver cells of people who do not drink alcohol excessively.

**Non-alcoholic steatohepatitis** A more advanced form of non-alcoholic fatty liver disease involving inflammation in and around the fatty liver cells. It may cause scarring of the liver and lead to cirrhosis.

**Obese** Having a body mass index $\geq 30$.

**Oesophageal varices** Varicose veins in the lower end of the oesophagus which develop as a complication of cirrhosis.

***p*-value** The probability that an observed difference could have occurred by chance if the null hypothesis is correct. A $p$-value of $< 0.05$ is conventionally considered to be statistically significant.

**Percutaneous** Performed through a needle puncture of the skin.

**Peripheral oedema** Swelling of the feet, ankles and legs.

**Portal hypertension** High blood pressure in the hepatic portal vein and its tributaries. It is often defined as a portal pressure gradient (the difference in pressure between the portal vein and the hepatic veins) of ≥ 5 mmHg. It may occur in people with or without chronic liver disease.

**Sclerotherapy** Injection of a medication into blood vessels or blood vessel malformations to make them shrink.

**Serum** The liquid component of blood (i.e. without the blood cells and clotting factors).

**Severe alcoholic hepatitis** Alcoholic hepatitis with symptoms severe enough to require hospital admission.

**Shear wave** An elastic wave which moves through the body of an object rather than over its surface.

**Spontaneous bacterial peritonitis** Infection of the fluid which can accumulate in the abdomen (ascites) in patients with severe liver disease.

**Steatosis** A condition characterised by the accumulation of excess fat within the liver cells.

**Transjugular** Via the jugular vein in the neck.

**Unit of alcohol** Quantity of any alcoholic drink which corresponds to approximately 10 ml (8 g) of ethanol.

**Variceal bleeding** Gastrointestinal bleeding caused by rupture of the oesophageal/gastric varices.

**Varices** Varicose veins usually in the stomach and lower end of the oesophagus (gullet) which develop as a complication of cirrhosis.

# List of abbreviations

| | |
|---|---|
| AAH | acute alcoholic hepatitis |
| ALD | alcohol-related liver disease |
| AUDIT | Alcohol Use Disorders Identification Test |
| ASH | alcoholic steatohepatitis |
| AUROC | area under the receiver operating characteristic curve |
| BMI | body mass index |
| CI | confidence interval |
| ELF test | Enhanced Liver Fibrosis test |
| FN | false-negative |
| FP | false-positive |
| HA | hyaluronic acid |
| HCC | hepatocellular cancer |
| HE | hepatic encephalopathy |
| HRAR | high-risk alcoholism relapse |
| HVPG | hepatic venous pressure gradient |
| ICU | intensive care unit |
| LSM | liver stiffness measurement |
| NAFLD | non-alcoholic fatty liver disease |
| NICE | National Institute for Health and Clinical Excellence |
| NILT | non-invasive liver test |
| PHt | portal hypertension |
| PICO | Population, Intervention, Comparator, Outcome |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| QALY | quality-adjusted life-year |
| QUADAS | QUality Assessment of Diagnostic Accuracy Studies |
| QUOROM | QUality Of Reporting Of Meta-analyses |
| RCT | randomised controlled trial |
| SGOT | serum glutamic oxaloacetic transaminase |
| TIMP-1 | tissue inhibitor of matrix metalloproteinase-1 |
| TN | true-negative |
| TP | true-positive |

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

# Executive summary

## Background

Excessive alcohol consumption may lead to the development of alcohol-related liver disease (ALD). ALD comprises a spectrum of disease, including hepatic steatosis (alcoholic fatty liver), alcoholic hepatitis, alcoholic fibrosis and cirrhosis, and hepatocellular cancer. In 2008, 0.95% of all deaths registered in people aged ≥ 20 years in England and Wales were attributed to ALD. Liver biopsy may be used in patients with suspected ALD to confirm the diagnosis, exclude other or additional liver pathologies, and provide accurate staging of the degree of liver injury in order to enable the prediction of prognosis and inform treatment decisions. However, as it is an invasive procedure that carries the risk of morbidity and mortality, current UK guidance recommends that biopsy is not required to confirm the diagnosis in patients with a high clinical suspicion of ALD in whom blood tests have excluded other causes of liver disease, unless it is necessary to confirm a diagnosis of acute alcoholic hepatitis in order to inform specific treatment decisions.

## Objectives

The objectives of this assessment are to evaluate the diagnostic accuracy, cost-effectiveness, and effect on patient outcomes of four non-invasive tests for liver fibrosis [the Enhanced Liver Fibrosis (ELF™) test (Siemens Healthcare Diagnostic Inc., Tarrytown, NY, USA), FibroTest (BioPredictive, Paris, France), FibroMAX (BioPredictive, Paris, France) and transient elastography (FibroScan®; produced by EchoSens, Paris, France and distributed in the UK by Artemis Medical Ltd, Kent, UK)] in patients suspected of having liver fibrosis related to alcohol consumption. The tests are assessed first as a replacement for liver biopsy, and secondly as an additional test prior to liver biopsy.

## Methods

A systematic review was undertaken to identify studies reporting the diagnostic and prognostic accuracy of the ELF test, FibroTest, FibroMAX and FibroScan for the identification of liver fibrosis and associated conditions in patients with suspected ALD. The following databases were searched in January 2010: MEDLINE (from 1950 to January 2010), MEDLINE In-Process & Other Non-Indexed Citations (from 1950 to January 2010), EMBASE (from 1980 to January 2010), Cochrane Database of Systematic Reviews (from 1996 to January 2010), Cochrane Central Register of Controlled Trials (from 1898 to January 2010), Cochrane Methodology Register (from 1904 to January 2010), Database of Abstracts of Reviews of Effects (from 1995 to January 2010), HTA Database (from 1995 to January 2010), NHS Economic Evaluation Database (from 1995 to January 2010), Cumulative Index to Nursing and Allied Health Literature (from 1982 to January 2010), Web of Knowledge, Science Citation Index, Conference Proceedings Citation Index, and BIOSIS Previews (from 1969 to January 2010). Research registers and conference proceedings were also searched. Study quality was assessed using the QUADAS (QUality Assessment of Diagnostic Accuracy Studies) checklist.

A mathematical model was constructed to estimate the incremental costs and incremental quality-adjusted life-years (QALYs) associated with the introduction of alternative strategies

compared with a biopsy-all strategy. Owing to the wide uncertainty in the data to populate key variables, 36 scenarios were assessed that varied the sensitivity of biopsy, the anxiety associated with biopsy, different values for the sensitivity and specificity for each non-invasive tests, and whether a percutaneous or transjugular biopsy was required. For each of these scenarios, nine strategies were evaluated, which were divided into triage strategies (where a positive test was confirmed by biopsy) and replacement strategies (where no confirmatory biopsy was provided). For each scenario and strategy, two threshold levels were reported where biopsying all patients was more cost-effective than the strategy: the decreased level of abstinence associated with the strategy compared with biopsying all and the level of QALY gain that would be required for a biopsy.

## Results

### *Summary of clinical results*

#### Diagnostic accuracy of the Enhanced Liver Fibrosis Test

No studies were identified that specifically assessed the ELF test. One study evaluated the diagnostic accuracy of the European Liver Fibrosis Test (essentially, the ELF test with the addition of age to the algorithm) compared with liver biopsy in patients with chronic liver disease, only 64 of whom had ALD; a follow-up study in 85 patients with ALD assessed its ability to predict long-term survival and relevant clinical events. This limited evidence suggests that, using a threshold score of 0.431, the European Liver Fibrosis Test can differentiate between moderate/severe fibrosis and milder/no fibrosis in patients with ALD with a sensitivity of 93% and a specificity of 100%; it is less good at identifying cirrhosis. It appears to have some predictive value in relation to both liver-related clinical outcomes and all-cause mortality. However, because the results rest on data from so few patients, evidence for the diagnostic and prognostic accuracy of the test is not robust.

#### Diagnostic accuracy of FibroTest

Five studies of FibroTest were identified. Two evaluated diagnostic test accuracy compared with liver biopsy in patients with known or suspected ALD. A further three recruited patients with liver disease of mixed aetiology, including ALD: the first assessed FibroTest's ability to identify portal hypertension (PHt) and also compared it with liver biopsy, the second assessed its ability to predict the presence of oesophageal varices, and the third assessed its predictive value in relation to survival at 2 and 6 months in patients with severe cirrhosis. Results from the largest study, which was also the most representative of the spectrum of patients with suspected ALD, suggest that, in such patients, using a threshold score of 0.30, FibroTest can differentiate between moderate/severe fibrosis and milder/no fibrosis with a sensitivity of 84% and specificity of 66%, while using a threshold score of 0.70, it can distinguish cirrhosis with a sensitivity of 91% and specificity of 87%. Very small studies suggest that, using a threshold score of 0.58, FibroTest can distinguish between patients with and without clinically significant PHt with a sensitivity of 93% and specificity of 87%, while, using a threshold score of 0.85, it can distinguish between those with and without grade 2 oesophageal varices with a sensitivity of 89% and specificity of 50%. However, the results relating to PHt and oesophageal varices are not robust because the studies were very small, and the conditions of interest were over-represented. FibroTest appears to predict survival with relatively low accuracy.

#### Diagnostic accuracy of FibroMAX

No relevant studies of FibroMAX were identified.

### Diagnostic accuracy of FibroScan

Six studies were identified that assessed the diagnostic test accuracy of FibroScan relative to liver biopsy in patients with known or suspected ALD. A further three studies recruited patients with liver disease of mixed aetiology, including ALD. One assessed the ability of FibroScan to predict the presence of large oesophageal varices in patients with cirrhosis, whereas the other two assessed its ability to predict clinically significant PHt. The study with the most representative population suggests that, using threshold scores of 5.9, 7.8, 11.0 and 19.5 kPa, respectively, FibroScan can differentiate between patients with and without fibrosis with a sensitivity of 83% and a specificity of 86%, and can identify moderate/severe fibrosis with a sensitivity of 80% and a specificity of 90.5%, severe fibrosis with a sensitivity of 87% and specificity of 80.5%, and cirrhosis with a sensitivity of 86% and specificity of 84%. However, again, these results are not robust because the study was relatively small and the conditions of interest were over-represented. FibroScan appears to be able to distinguish between patients with and without PHt, and with less success between patients with and without large oesophageal varices. There are no long-term data relating FibroScan results to survival or other clinical outcomes.

### Adverse effects and contraindications

The non-invasive tests included in this review appear to be safe. The adverse events associated with the ELF test, FibroTest, and FibroMAX are those associated with diagnostic venepuncture generally: primarily pain and bruising, with occasional vasovagal reactions and very rarely potentially disabling nerve injuries. There is no evidence to indicate that FibroScan is specifically associated with any adverse effects. By contrast, liver biopsy is associated with a high level of morbidity and occasional mortality.

No contraindications have been specified for the ELF test. The contraindications specified for FibroTest, FibroMAX, and FibroScan all relate to the mode of operation of the test, and do not relate to any potential for harm in patients with the relevant characteristics, although they will restrict their practical utility. The most important of these limitations is the restriction on the use of FibroScan in obese patients.

## Summary of cost-effectiveness and benefits versus risks

It was concluded that no robust estimate could be provided regarding the incremental costs, incremental QALYs and, therefore, the cost per QALY of a strategy. Scenarios exist in which each of the strategies analysed is more cost-effective than biopsying all patients and, in contrast, scenarios exist in which each strategy is less cost-effective than biopsying all patients. No conclusive result can be provided on the most cost-effective strategy until further data are available; however, there is evidence that some strategies, such as using clinical experience or diagnosing all patients with cirrhosis, will not be the most cost-effective.

## Conclusions

### *Implications for service provision*

Owing to the lack of a conclusion regarding the cost-effectiveness of the strategies, it is anticipated that there would be no change in service provision.

### *Suggested research priorities*

A large number of parameters require data; however, the following are selected as being of most importance:

- the sensitivity and specificity of liver biopsy against a gold-standard of post-mortem evaluation of fibrosis
- the sensitivity and specificity of each non-invasive liver test (NILT) against a gold standard of post-mortem evaluation of fibrosis (or failing this biopsy at validated and pre-selected cut-off thresholds for the various degrees of liver damage)
- the influence of potential confounding variables such as current drinking behaviour and the degree of hepatic inflammation on the performance of NILTs
- differential information on the percentage of alcohol misusers who will develop alcohol-related cirrhosis over time, by age at onset, gender and ethnic origin
- the likelihood, and magnitude, of decreases in abstinence rates associated with a diagnosis of significant ALD by diagnostic modality
- the incidental gains in QALYs that may be associated with biopsy, because of the determination of non-ALD-related aetiologies.

## Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

# Chapter 1

# Background

## Description of health problem

Alcohol-related liver disease (ALD) is the term used to describe changes within the liver that develop as a result of excess alcohol consumption. Included within this spectrum are hepatic steatosis (alcoholic fatty liver), alcoholic hepatitis, alcoholic cirrhosis, and hepatocellular cancer (HCC).[1–3] Although patients with cirrhosis may remain free of major complications for several years, once such complications develop survival is usually short.[4]

In 2008, 4764 deaths registered in England and Wales were attributed to ALD.[5]

### *Aetiology, pathology and prognosis*

#### Aetiology

Alcohol-related liver disease is caused by the consumption of substantial quantities of alcohol. In the UK, alcoholic drinks are measured in units corresponding to approximately 10 ml (8 g) of ethanol.[6] Previously, the minimum alcohol intake associated with significant liver damage was thought to be around 20 units a day for 5 years and the likelihood of light to moderate drinkers developing cirrhosis was considered very remote.[7] More recently, it has been suggested that all people whose daily alcohol consumption exceeds 80 g (10 units) will eventually develop steatosis; 10–35% will develop alcoholic hepatitis, and approximately 10% will develop cirrhosis.[8] However, because steatosis is usually, and hepatitis and cirrhosis sometimes, asymptomatic, many will be unaware that they have ALD; indeed, 30–40% of cases of cirrhosis are not discovered until autopsy.[1]

#### Pathology

As noted above, the term ALD covers a spectrum of illness. The majority of individuals who abuse alcohol will develop steatosis,[9] a condition characterised by the accumulation of excess fat within the liver cells.[10] Steatosis is usually asymptomatic, but may manifest as changes in liver function test results; it is reversible if alcohol consumption is stopped or significantly reduced.[1] However, approximately 20% of people with alcoholic steatosis who continue long-term alcohol consumption are likely to develop fibrosis and cirrhosis.[2]

In a minority (perhaps < 20%) of alcohol misusers, alcoholic steatosis evolves with the development of inflammation and cell death within the liver. This phase of injury is known as alcoholic hepatitis.[10] The majority of patients who develop alcoholic hepatitis are asymptomatic, but others suffer non-specific symptoms such as nausea, vomiting, and abdominal pain, whereas others present with acute alcoholic hepatitis (AAH), which is characterised by abdominal pain/discomfort, fever, a high white blood cell count, abnormalities of blood clotting, jaundice, and other features of liver failure.[1] Patients with mild alcoholic hepatitis will recover well, with reversal of the liver damage, if they stop drinking. Progression to cirrhosis is more likely in women than in men, and in individuals of both sexes who have severe alcoholic hepatitis on presentation, irrespective of their future drinking behaviour.

Repeated episodes of alcoholic hepatitis may result in the development of scar tissue known as fibrosis; this can be localised or diffuse, and can eventually lead to distortion of the architecture of the liver by broad fibrous bands and the development of regenerative nodules, which are the hallmark of cirrhosis. Patients with cirrhosis may be asymptomatic, and their liver function tests may show few if any abnormalities; such patients are described as having compensated disease.[10,11] In a recent Danish study, 24% of patients diagnosed with alcoholic cirrhosis had no complications.[12] Patients with cirrhosis may, however, develop a number of complications as a result of hepatocellular failure and the development of portal hypertension (PHt). Hepatocellular failure results in loss of the liver's ability to deal with waste products (e.g. bilirubin from red blood cells, toxins produced in the bowel such as ammonia, and drugs) and impairment of its synthetic function leading to low levels of body proteins such as albumin and the factors responsible for blood clotting. PHt develops when the passage of blood from the intestine and the spleen, which is usually cleansed and detoxified in the liver, is impeded because of the presence of fibrosis. As a result, collateral vessels develop to bypass the liver to ensure that the blood gets back to the heart. These collateral vessels may develop in the stomach and oesophagus and may rupture when the pressure gradient within the portal system [the hepatic venous pressure gradient (HVPG)] exceeds 12 mmHg.[13] Hepatocellular failure and PHt result in the development of several problems, including jaundice, ascites, variceal bleeding, and hepatic encephalopathy (HE).[4] Patients with these complications are described as having decompensated cirrhosis. Each year, approximately 5–7% of patients with compensated cirrhosis will develop decompensated disease.[4] Approximately 20% of patients with cirrhosis will go on to develop HCC; this is more common in men than in women and in those who have stopped drinking (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication).

Alcoholic hepatitis and cirrhosis may co-exist within the same patient: thus, it has been reported that 50–60% of patients diagnosed with AAH already have cirrhosis at the time of diagnosis.[14]

## Measurement of disease

Traditionally, liver biopsy has been used to obtain histological samples from patients suspected of having ALD. These biopsy samples are used to confirm the diagnosis of ALD, exclude other or additional liver pathologies, and provide accurate staging of the degree of liver injury in order to enable the prediction of prognosis and inform treatment decisions.

Various semi-quantitative scoring systems have been developed to measure and categorise disease progression in biopsy samples. The most widely used system, the METAVIR system, has a grading system that scores the degree of inflammation from 0 to 4 (where 0 is no activity and 4 is severe activity) and a staging system that categorises the degree of architectural change and consequent severity of the underlying liver disease[15] from none (F0) to mild (F1), moderate (F2), severe (F3), and cirrhosis (F4). It should be noted that METAVIR, and other similar systems, have been described as being, at best, ordered categorical data;[16] they do not represent an arithmetic progression, and consequently, for example, a METAVIR score of F4 does not involve twice as much scarring as a score of F2. For further details of the various scoring systems referred to in this review, see *Appendix 1*.

In clinical practice, various boundaries are said to be significant. These are the boundaries between no-to-mild fibrosis (F0–F1) and significant fibrosis ($F \geq 2$),[17] and between mild-to-moderate fibrosis (F1–F2) and severe fibrosis or cirrhosis (F3–F4).[18] However, the former seems primarily relevant in viral hepatitis, where F2 fibrosis forms the threshold for the initiation of interferon therapy;[19] it arguably has less relevance in ALD, where the boundary between F0–F3 and F4 appears more important as it forms the threshold for the initiation of regular screening for varices and HCC.

## Prognosis

Prognosis in patients with ALD is strongly linked to the degree of disease progression. In an early UK study, Saunders *et al.*[20] found that 5-year survival in adults identified between 1959 and 1976 as having alcoholic cirrhosis was only 36%; most already had decompensated cirrhosis at diagnosis. More recently, Jepsen *et al.*[12] found that the median survival in Danish patients diagnosed between 1993 and 2005 with alcoholic cirrhosis without complications was 48 months. The 1-year mortality ranged from 17% in those who presented with no complications to 64% in those who presented with HE; the equivalent 5-year mortality figures were 58% and 85%, respectively. More generally, median survival times for patients diagnosed with compensated and decompensated cirrhosis have been estimated at over 12 years and around 2 years, respectively; most patients originally diagnosed with compensated cirrhosis die only after the development of decompensation.[4]

In patients with cirrhosis, the presence of superimposed alcoholic hepatitis and the development of hepatocellular carcinoma significantly reduced survival. Verrill *et al.*[21] found that 30-day mortality in patients with cirrhosis and alcoholic hepatitis was 27%, while Orrego *et al.*[22] found that the 1-year and 5-year mortality were significantly higher than in those of patients who had alcoholic cirrhosis alone, and Chedid *et al.*[23] found that 65% of patients with concurrent alcoholic hepatitis and cirrhosis died within 4 years of diagnosis, mostly within the first year. Alcoholic cirrhosis substantially increases the risk of hepatocellular carcinoma,[1] which, again, has a poor prognosis: 1-year survival after diagnosis of HCC is about 20%, 5-year survival about 5%,[24] and approximately 15% of patients who develop alcohol-related cirrhosis will die of HCC.[9]

However, the outcome for patients with alcohol-related cirrhosis is substantially determined not only by the degree of decompensation at presentation but also by subsequent drinking behaviour. In 2004, Tome and Lucey[25] suggested that, in people with clinically compensated cirrhosis, the 5-year survival was about 90% in those who abstained compared with about 70% in those who did not, whereas in 2003 Mann *et al.*[1] suggested that, in patients with late-stage, decompensated cirrhosis, the 5-year survival was 60% in those who abstained, but only 35% in those who did not. In a recent study, Verrill *et al.*[21] found that patients who reported being abstinent from alcohol 30 days after receiving a diagnosis of alcoholic cirrhosis had a 7-year survival of 72% compared with 44% in those who continued to drink. A considerably older study by Saunders *et al.*[20] found that 10-year survival in patients who abstained completely from alcohol was nearly 60% in those who presented with compensated cirrhosis and around 50% in those who presented with decompensated cirrhosis, compared with around 30% and < 10%, respectively, in those who continued drinking. It is currently estimated that a middle-aged man who presents with compensated cirrhosis and subsequently abstains from alcohol has a 60–80% chance of being alive in 10 years, whereas a similar individual who presents with, for example, variceal bleeding, and who survives the initial presentation but who continues to drink, is unlikely to survive more than a year or two (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication).

A small study by Day,[26] of 96 patients with biopsy-proven ALD and a weekly alcohol intake of 135 ± 8 units, suggests that a reduction in alcohol intake may also be beneficial. No difference was found between abstainers and mild/moderate drinkers in either mortality (total or liver-related) or the development of new liver complications; however, continued heavy drinking (> 50 units/week for men and > 35 units/week for women) was predictive of both total and liver-related mortality.

Alcohol consumption also has a marked impact on prognosis in patients with milder forms of ALD. With abstinence, hepatic steatosis appears to be completely reversed within several weeks and AAH usually improves.[25] Patients with hepatic steatosis who stop drinking have a near-normal

lifespan; future drinking behaviour is the most important determinant of outcome in this patient group. Their prognosis has improved over time as the treatment of their complications has become more successful. Moreover, as Chedid *et al.* have shown,[23] these patients tend to die as a result of their alcoholism rather than their liver disease. Thus, although the 2-year survival in patients with alcoholic steatosis is only 70%, those who die do so as a result of other alcohol-related events such as accidents, injuries, suicide, and homicide, not their liver disease. However, although some patients with alcoholic hepatitis who abstain from alcohol make a complete recovery, others develop cirrhosis despite abstinence.[10]

Although only a minority of patients with ALD abstain from alcohol following diagnosis, the proportion may be increasing; although only 19% of patients in a UK study who were diagnosed with alcoholic cirrhosis between 1959 and 1976 abstained from alcohol after hospital discharge,[20] 36% of patients in a Danish study who received a hospital diagnosis of cirrhosis between 1993 and 2005 maintained abstinence,[12] and 46% of participants in a UK study who were diagnosed with biopsy-proven alcohol-related liver cirrhosis between 1995 and 2000 reported abstinence both at 30 days post diagnosis and at follow-up a mean of 3.4 years later (although some admitted to lapses during that period).[21] However, the precise diagnosis given may influence subsequent drinking behaviour. Thus, Day[26] studied 96 patients with biopsy-proven ALD referred to unit between August 1991 and August 1993: 76% had established cirrhosis or fibrosis, 23% had alcoholic hepatitis and 8% steatosis. At follow-up, 50% were still drinking heavily, although 21% were completely abstinent, and 21% had mild and 8% moderate alcohol intake. The only predictor of continued heavy intake was the absence of cirrhosis; continued heavy drinking was reported in 59% of patients without, and 38% of patients with, cirrhosis ($p = 0.04$). By contrast, a study of patients receiving liver transplants for ALD found that a number of factors were independently associated with a significantly increased risk of resumption of harmful alcohol consumption after transplantation: these were pre-transplant diagnosis of an anxiety or depressive disorder, a pre-transplant period of abstinence < 6 months, and a total score of > 3 on the high-risk alcoholism relapse (HRAR) scale (*Table 1*). Over a mean follow-up period of 61.2 months, the overall relapse rate was 11.9%; however, it varied from 5% to 100% depending on the number of relevant factors (*Table 2*).[27]

It has therefore been suggested that, in patients with ALD, the degree of fibrosis may have less prognostic importance than other factors such as the severity of alcoholic hepatitis, severity of

**TABLE 1** High-risk alcoholism relapse (HRAR) scale.[27] Reproduced with permission from De Gottardi A, Spahr L, Gelez P, Morard I, Mentha G, Guillaud O, *et al*. A simple score for predicting alcohol relapse after liver transplantation: results from 387 patients over 15 years. *Arch Intern Med* 2007;**167**:1183–8.

| Item | Score |
|---|---|
| ***Duration of heavy drinking, in years*** | |
| ≤ 11 | 0 |
| 11–25 | 1 |
| ≥ 25 | 2 |
| ***Number of daily drinks (one drink = 12 g of ethanol)*** | |
| ≤ 9 | 0 |
| 9–17 | 1 |
| ≥ 17 | 2 |
| ***Number of prior alcoholism inpatient treatments*** | |
| 0 | 0 |
| 1 | 1 |
| ≥ 1 | 2 |

**TABLE 2** Resumption of harmful drinking following liver transplant for ALD[27]

| Number of relevant factors[a] | Number resuming harmful alcohol consumption | Rate % (95% CI) | Mean time to relapse (months) |
|---|---|---|---|
| 0 | 13/272 | 5 (*2 to 7*) | 45 |
| 1 | 16/92 | *17 (10 to 25)* | 30 |
| 2 | 14/22 | 64 *(44 to 84)* | 32 |
| 3 | 3/3 | 100 | 23 |

CI, confidence interval.
a Pre-transplant diagnosis of anxiety or depressive disorder, pre-transplant period of abstinence < 6 months, or total score > 3 on the HRAR scale.
Data in roman were taken directly from the text, data in *italics* were calculated by the reviewers.

liver dysfunction as assessed by tools such as the Child–Pugh score, the MELD (Model for End-Stage Liver Disease) score, or the Glasgow score and subsequent drinking behaviour.[28]

## *Epidemiology demographic factors (age, sex, ethnicity, income, and regional variation)*

Current government guidelines state that women should drink no more than 2–3 units of alcohol a day, and men no more than 3–4 units; the recommended weekly limits are 14 and 21 units, respectively.[29] However, in 2007, a health survey for England found that 31% of women and 42% of men in England aged ≥ 16 years reported drinking more than the government recommended maximum on at least 1 day in the preceding week.[30,31] The figures reported to the Welsh health survey in 2008 were higher, at 38% of women and 52% of men.[32]

Hazardous drinking has been defined as an established pattern of drinking that carries a risk of physical or psychological harm. Harmful drinking has been defined as drinking at a level where damage to physical or mental health is considered likely.[31] A score of 8–15 on the Alcohol Use Disorders Identification Test (AUDIT) is held to represent hazardous drinking, and an AUDIT score of ≥ 16 as harmful drinking (for the AUDIT tool, see *Appendix 2*).[31] A total of 20.4% of individuals in England aged ≥ 16 years living in private households who participated in the 2007 Adult Psychiatric Morbidity Survey reported established patterns of alcohol consumption that were categorised as hazardous, and a further 3.8% reported patterns categorised as harmful (*Table 3*).[31] These figures are probably underestimates as they are likely to under-represent both alcohol-dependent individuals, who are more likely than non-alcohol-dependent individuals to be homeless or living in an institution, and problem drinkers living in private households, who may be less likely than those who are not problem drinkers to participate in surveys.[31]

In England, rates of hazardous or harmful drinking vary by ethnic group, being highest in white people and lowest in people of South Asian origin[31] (*Table 4*). However, there is evidence both that alcohol problems may be increasing in the South Asian community in the UK, and that they may be particularly vulnerable to ALD.[33,34]

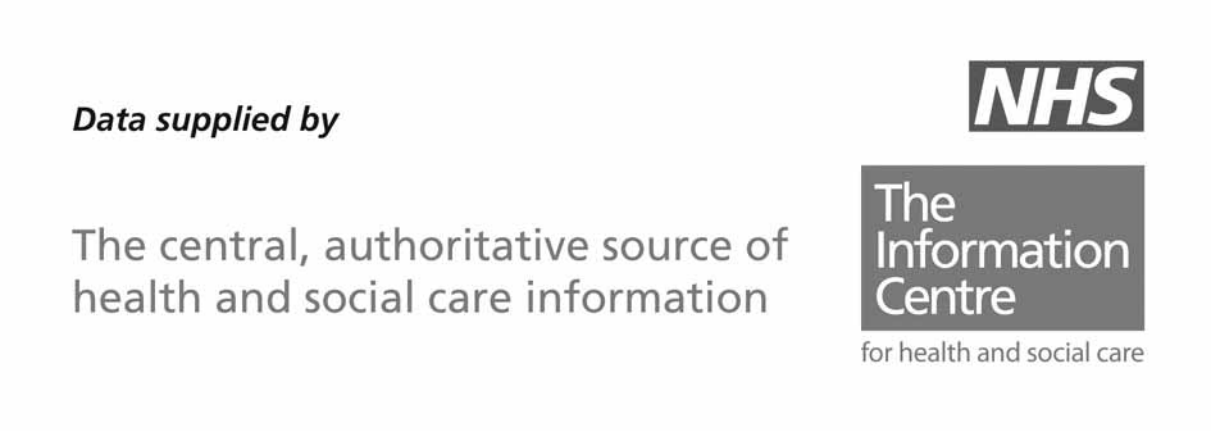
In England, levels of hazardous or harmful drinking vary by region. In 2007, they were lowest for women in the East of England (12.2%) and highest in Yorkshire and the Humber (21.1%), whereas they were lowest for men in the East Midlands (27.8%) and highest in the North East (42.4%) (*Table 5*).[31]

Data about levels of hazardous or harmful drinking are not available for Wales. Although, in 2008, 45% of all adults in Wales reported drinking above the government-recommended sensible levels on at least 1 day in the preceding week,[32] not all of these would necessarily have drinking patterns that equate to an AUDIT score of ≥ 8.

**TABLE 3** Prevalence of hazardous and harmful drinking in people aged ≥16 years in England, 2007, by sex (from Fuller *et al*.[31])

| AUDIT score | Men (%) | Women (%) | All (%) |
|---|---|---|---|
| 0–7: not hazardous | 66.8 | 84.3 | 75.8 |
| 8–15: hazardous, not harmful | 27.4 | 13.8 | 20.4 |
| 16–40: harmful | 5.8 | 1.9 | 3.8 |
| 8–40: hazardous or harmful | 33.2 | 15.7 | 24.2 |

**TABLE 4** Prevalence of hazardous and harmful drinking in people aged ≥16 years (age-standardised) in England, 2007, by ethnicity and sex (from Fuller *et al*.[31]) 




| | Ethnicity (%) | | | |
|---|---|---|---|---|
| AUDIT score | White | Black | South Asian | Other[a] |
| ***Men*** | | | | |
| 0–7: not hazardous | 64.2 | 81.4 | 88.0 | 84.1 |
| 8–15: hazardous, not harmful | 29.6 | 15.6 | 9.9 | 13.8 |
| 16–40: harmful | 6.2 | 3.0 | 2.1 | 2.1 |
| 8–40: hazardous or harmful | 35.8 | 18.6 | 12.0 | 15.9 |
| ***Women*** | | | | |
| 0–7: not hazardous | 83.4 | 95.4 | 96.9 | 84.5 |
| 8–15: hazardous, not harmful | 14.5 | 4.6 | 3.1 | 13.9 |
| 16–40: harmful | 2.0 | 0 | 0 | 1.6 |
| 8–40: hazardous or harmful | 16.6 | 4.6 | 3.1 | 15.5 |

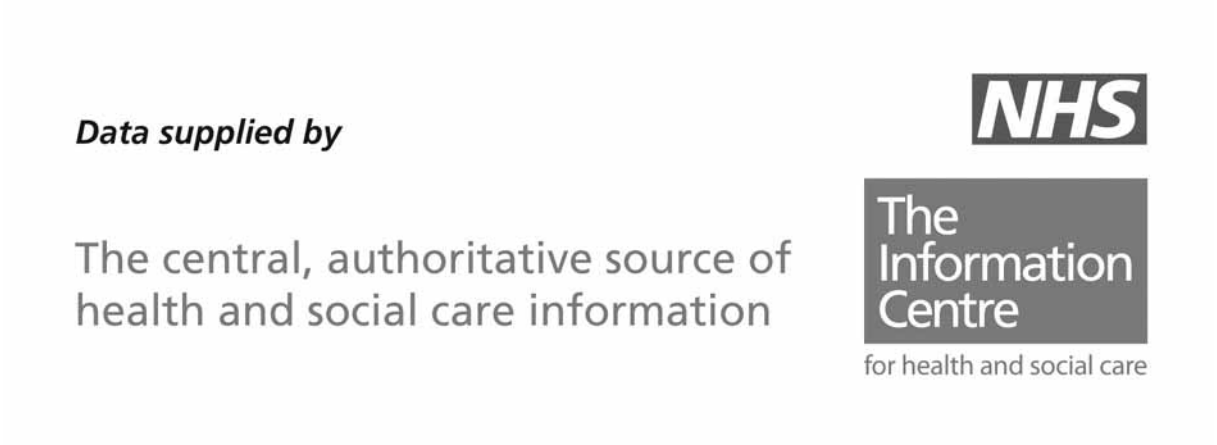
a Includes Chinese and mixed ethnic groups.

## *Incidence and prevalence*

Alcohol-related liver disease is a major cause of death in England and Wales. In 2008, 4764 deaths registered in England and Wales were attributed to ALD[5] – 0.95% of all deaths registered in people aged ≥20 years.

Although ALD mortality in England and Wales appears to have risen by 450% over the past 30 years,[35] this rise in registrations may be at least in part because of the under-reporting of alcohol-related deaths during the earlier part of that period. This may be explained by the then

**TABLE 5** Prevalence of hazardous and harmful drinking in people aged 16 and over (age-standardised) in England, 2007, by region and sex (from Fuller *et al*.[31]) Copyright© 2011, re-used with the permission of The Health and Social Care Information Centre. All rights reserved.




| AUDIT score | Government Office region (%) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | North East | North West | Yorkshire and the Humber | East Midlands | West Midlands | East of England | London | South West | South East |
| ***Men*** | | | | | | | | | |
| 0–7: not hazardous | 57.6 | 61.9 | 59.4 | 72.2 | 67.0 | 65.9 | 70.0 | 68.9 | 71.2 |
| 8–15: hazardous, not harmful | 32.2 | 31.7 | 34.4 | 23.8 | 26.2 | 27.9 | 25.6 | 25.5 | 23.4 |
| 16–40: harmful | 10.2 | 6.4 | 6.2 | 4.0 | 6.7 | 6.2 | 4.4 | 5.7 | 5.4 |
| 8–40: hazardous or harmful | 42.4 | 38.1 | 40.6 | 27.8 | 33.0 | 34.1 | 30.0 | 31.1 | 28.8 |
| ***Women*** | | | | | | | | | |
| 0–7: not hazardous | 79.2 | 80.6 | 78.9 | 82.9 | 84.5 | 87.8 | 86.2 | 85.7 | 87.7 |
| 8–15: hazardous, not harmful | 17.0 | 17.1 | 18.4 | 15.0 | 13.5 | 11.5 | 12.5 | 12.7 | 10.4 |
| 16–40: harmful | 3.7 | 2.3 | 2.7 | 2.2 | 2.0 | 0.6 | 1.3 | 1.7 | 1.9 |
| 8–40: hazardous or harmful | 20.8 | 19.4 | 21.1 | 17.1 | 15.5 | 12.2 | 13.8 | 14.3 | 12.3 |

requirement for post-mortem examinations and coroners' inquests following alcohol-related deaths, and the potential invalidation of life insurance (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication). It may also be due in part to the increases in obesity and type 2 diabetes observed during the later part of the period. In a prospective study in France, Raynard *et al*.[36] found that body mass index (BMI) and fasting blood glucose were risk factors for histologically confirmed fibrosis in ALD, independent of age, gender, daily alcohol intake, and total duration of alcohol misuse, whereas, in Scotland, two prospective cohort studies found that raised BMI and alcohol consumption had a supra-additive effect on liver disease mortality.[37] The

proportion of adults in England with a BMI ≥ 30 (i.e. categorised as obese or morbidly obese) rose from 13% of men and 18% of women in 1993, the first year for which data are available, to 25% and 28%, respectively, in 2008, whereas the proportions with raised waist circumference (i.e. > 102 cm for men and > 88 cm for women) rose from 20% of men and 26% of women in 1993 to 34% and 44%, respectively, in 2008,[38] and, between 1991 and 2006, the prevalence of diagnosed type 2 diabetes increased by 65% in men and by 25% in women.[39]

In 2008, over twice as many deaths were attributed to ALD in men than in women (*Table 6*).

Between 1992 and 2001, 3360 incident cases of cirrhosis in patients aged ≥ 25 years were reported to the UK General Practice Research Database, which contains data relating to over 13 million patients in the UK[40] (approximately 20% of the UK population). Of the reported cases, 1287 (38.3%) were judged, on the basis of records of problem drinking in the GP notes, to be alcoholic cirrhosis.[40] These figures suggest an annual incidence of alcoholic cirrhosis in the UK of approximately 25,740. During the 10 years from 1992 to 2001, the incidence of alcoholic cirrhosis rose by 75% in men and by 34% in women,[40] and in 2000 it was claimed that around 80% of all cases of liver cirrhosis seen in district general hospitals in the UK could be attributed to alcohol.[8]

## *Impact of health problem*

### Significance for patients in terms of ill-health (burden of disease)

Alcohol-related liver disease is often asymptomatic. However, cirrhosis is associated with a number of complications, the more common of which include variceal bleeding, ascites, spontaneous bacterial peritonitis, and encephalopathy;[41] these are associated with a significant burden of disease. Moreover, approximately 20% of patients with cirrhosis will go on to develop HCC (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication).

**TABLE 6** Deaths attributed to ALD (ICD-10 code K70), 2008, by age and gender[5]

| | Male | | Female | |
|---|---|---|---|---|
| Age group (years) | Number of deaths | Percentage of total deaths | Number of deaths | Percentage of total deaths |
| All ages | 3200 | 1.3 | 1564 | 0.6 |
| 20–24 | 2 | 0.2 | 5 | 1.1 |
| 25–29 | 25 | 1.8 | 10 | 1.7 |
| 30–34 | 83 | 4.9 | 36 | 4.2 |
| 35–39 | 173 | 6.5 | 85 | 6.0 |
| 40–44 | 339 | 9.0 | 172 | 7.2 |
| 45–49 | 408 | 7.8 | 217 | 6.4 |
| 50–54 | 562 | 8.0 | 252 | 5.2 |
| 55–59 | 534 | 5.0 | 256 | 3.6 |
| 60–64 | 471 | 2.9 | 225 | 2.0 |
| 65–69 | 289 | 1.4 | 143 | 1.1 |
| 70–74 | 176 | 0.6 | 105 | 0.5 |
| 75–79 | 89 | 0.2 | 41 | 0.1 |
| 80–84 | 32 | 0.07 | 10 | 0.02 |
| 85–89 | 14 | 0.04 | 7 | 0.01 |
| 90–94 | 3 | 0.02 | 0 | 0.0 |
| ≥ 95 | 0 | 0.0 | 0 | 0.0 |

ICD–10, *International Classification of Diseases*,10th Edition.

Portal hypertension is a major complication of chronic liver disease. It contributes to the development of ascites and HE, and forms a direct cause of variceal bleeding. Clinically significant PHt (i.e. PHt associated with a risk of such complications) has been defined as a HVPG measurement above about 10 mmHg.[42] Patients with clinically significant PHt should be offered treatment to reduce the risk of complications.[42]

Approximately 40% of patients with liver cirrhosis have oesophageal varices, and approximately one-third of these will suffer variceal bleeding within 2 years of diagnosis. Such bleeding is associated with a mortality rate of 20–40% per episode[43] and a 1-year mortality of 57%;[4] nearly half of the deaths occur within 6 weeks of the initial episode of bleeding.[4]

Over a 10-year period, approximately 50% of patients with compensated cirrhosis will develop ascites (excessive fluid within the peritoneal cavity).[44] This is associated with a poor quality of life, increased risk of infections, renal failure, and poor long-term outcomes.[45]

Cirrhosis is also associated with HE, a condition that encompasses mental alterations ranging from mild (trivial lack of awareness, a shortened attention span, or euphoria or anxiety) to more obvious mental alterations including lethargy or apathy, occasional disorientation, personality change, and inappropriate behaviour, and in more severe cases to somnolence, confusion, gross disorientation and bizarre behaviour, and ultimately coma. Motor function is also impaired.[46] Patients who do not display any overt neurological abnormalities may nonetheless display subtle abnormalities of cognition and/or neurophysiological variables; this condition is termed minimal HE. Minimal HE appears to be associated with a reduction in health-related quality of life and in the ability to perform complex tasks such as driving a car.[46] It has been suggested that approximately 35–40% of cirrhotic patients will develop overt HE at some point, whereas approximately 20–60% with liver disease will develop minimal HE.[47]

## Significance for the NHS

The number of patients admitted to NHS hospitals in England with ALD has risen year on year from around 25,700 in 2002–3 to approximately 38,300 in 2007–8, an increase of around 49%. In approximately 14,300 of these patients (37%), ALD was the primary diagnosis (*Table 7*).[48]

In 2006, liver disease was responsible for approximately 1600 hospital admissions in Wales.[49] No information was provided regarding the proportion of these admissions that could be attributed specifically to ALD.

**TABLE 7** Alcohol-related liver disease (ICD-10 code K70): hospital admissions in England (after *NHS Information Centre Statistics on Alcohol: England, 2009*, Tables 4.3 and 4.7[48])

| | Number of admissions, rounded to nearest hundred | | | | | |
|---|---|---|---|---|---|---|
| Hospital admissions related to ALD | 2002–3 | 2003–4 | 2004–5 | 2005–6 | 2006–7 | 2007–8 |
| Patients admitted with ALD | 25,700 | 28,600 | 31,500 | 34,400 | 37,700 | 38,300 |
| Patients admitted primarily because of ALD | 11,500 | 12,200 | 13,100 | 13,800 | 14,500 | 14,300 |
| Alcoholic fatty liver | 100 | 200 | 200 | 200 | 200 | 200 |
| Alcoholic hepatitis | 1100 | 1200 | 1200 | 1300 | 1400 | 1400 |
| Alcoholic fibrosis and sclerosis of liver | 100 | 100 | 100 | 100 | 100 | 100 |
| Alcoholic cirrhosis of liver | 3100 | 3400 | 3800 | 4200 | 4800 | 4800 |
| Alcoholic hepatic failure | 800 | 800 | 900 | 1000 | 1100 | 1100 |
| Alcoholic liver disease, unspecified | 6300 | 6500 | 6800 | 7000 | 7000 | 6700 |

ICD–10, *International Classification of Diseases*, 10th Edition..

The number of patients admitted to adult, general critical/intensive care units (ICUs) in England and Wales with ALD is estimated to have increased from 550 in 1996 to 1513 in 2005, and the number of ICU bed-days which they occupied to have increased from around 3100 to > 10,000. These figures are likely to be underestimates, as they exclude admissions to specialist liver critical-care beds and also any ICU patients with ALD who did not have ALD recorded as a primary or secondary reason for admission.[50]

## Current service provision

### *Diagnosis and subsequent management of disease*

#### Diagnosis

The diagnostic pathways for suspected ALD are complex. Many people with ALD have no signs or symptoms of disease, the first indication of liver disease often coming from routine liver function tests. Others first come to medical attention when they report relatively mild symptoms (e.g. nausea, vomiting, abdominal discomfort, or diarrhoea). Sometimes, ALD is identified when people present voluntarily for detoxification, or when they require treatment for alcohol-related injuries or pneumonia, or alcoholic damage to organs other than the liver (e.g. the pancreas, heart, brain or peripheral nerves).[8] Yet other patients do not present until they have advanced liver disease characterised by more severe symptoms, such as jaundice, ascites, encephalopathy, or upper gastrointestinal bleeding.[8]

The variability in the diagnostic pathways, caused by the varying reasons for presentation, is complicated by the absence of a single test that differentiates ALD from liver disease of other aetiologies. Rather, the patient's history is obtained to identify risk factors (alcohol and other) for liver disease, and liver function tests, blood counts, and hepatitis serology are performed to exclude liver diseases of other aetiologies.[51] Ultrasound may also be used.[9] Because patients without clinical evidence of decompensated cirrhosis may have histologically advanced ALD but normal, or only mildly abnormal, liver function test results, liver biopsy may be required to confirm the diagnosis and the stage of the disease by assessing the degree of fibrosis.[9,52]

There is a lack of consensus about the role and the timing of biopsy in patients with suspected ALD.[51] This derives in part from the absence of high-quality evidence for the accuracy of liver biopsy specifically in the diagnosis of ALD, together with the fact that its status as the 'gold standard' diagnostic and staging tool has been called into question for several reasons. There are also concerns relating to the safety of liver biopsy. For further details, see *Summary of diagnostic tests*.

#### Management of disease

Following a diagnosis of ALD, the aims of management are to treat the underlying cause, prevent disease progression, and manage complications. By far the most important management aim in all patients is to ensure long-term abstinence from alcohol. As noted in *Aetiology, pathology and prognosis*, a small UK study found that ≤ 50% of patients with ALD either abstained completely from alcohol or significantly reduced their intake following simple advice from a physician during their initial presentation.[26] Additional pharmacological or non-pharmacological interventions may also be used to promote abstinence.[8,52] Other aspects of management include lifestyle changes to reduce cigarette consumption and obesity,[10] if relevant.

In patients with relatively mild alcoholic hepatitis, the focus of treatment is the achievement of abstinence. Corticosteroid therapy may be used to treat severe AAH,[9] but there is no conclusive evidence that it leads to significant improvements in survival.[53] As most patients with AAH have some degree of malnutrition, nutritional therapy may also be offered.[25]

A range of therapies may be used to treat the various complications of alcoholic cirrhosis.[8,52] Complications such as fluid retention, HE, and variceal bleeding are treated symptomatically. Because prophylactic treatment with non-selective beta-blockers or band ligation reduces the risk of first bleeding in patients with medium or large varices by 50%, all patients with cirrhosis should be screened regularly by endoscopy for signs of PHt (particularly the presence of oesophageal varices), and, if necessary, offered preventative treatment against gastrointestinal bleeding. Current guidance recommends that all patients diagnosed with cirrhosis should be offered such endoscopic screening every 2 years if they have no varices and annually if they have small varices[54] or decompensated disease with or without varices.[42] Abdominal Doppler ultrasound may be used to screen for liver and portal abnormalities (especially HCC).[11]

Orthotopic liver transplantation has a place in the management of patients with decompensated alcohol-related cirrhosis who have failed to improve despite well-documented abstinence from alcohol and expert medical treatment for a period of approximately 6 months. Survival rates are similar to those observed in patients transplanted for non-alcoholic cirrhosis, although recidivism rates are still unacceptably high in some centres. Transplantation is also considered the optimal treatment for early HCC.[55] However, the supply of donor livers is limited.[8]

### *Current service cost*

The current service cost has been limited to the cost of diagnostic liver biopsy, which has been estimated at £894 for a percutaneous biopsy[56] and £1500 for a transjugular biopsy.[9] It is recognised that there are associated costs that have not been estimated, but that are assumed to be independent of the strategy employed.

No evidence has been identified regarding the number of diagnostic liver biopsies undertaken each year in England and Wales specifically as a consequence of suspected ALD.

### *Relevant national guidelines, including national service frameworks*

The following relevant national guideline was issued in 2010:[9]

> Alcohol use disorders: diagnosis and clinical management of alcohol-related physical complications.
>
> The National Clinical Guideline Centre for Acute and Chronic Conditions

### *Variation in services and uncertainty about best practice*

It became apparent, during the course of this project, that rates of referral to secondary care of patients with suspected ALD vary considerably in different parts of the country and that there is some uncertainty about best practice.

## Description of technology under assessment

### *Summary of diagnostic tests*

This review assesses four non-invasive tests for liver fibrosis: three blood tests [the Enhanced Liver Fibrosis (ELF™) test (Siemens Healthcare Diagnostic Inc., Tarrytown, NY, USA), FibroTest (BioPredictive, Paris, France) and FibroMAX (BioPredictive, Paris, France)] and transient elastography (FibroScan®; produced by EchoSens, Paris, France and distributed in the UK by Artemis Medical Ltd, Kent, UK). The reference standard with which they are generally compared is liver biopsy, but HVPG measurement and upper gastrointestinal endoscopy have also been considered. All these tests are discussed in turn below.

## The Enhanced Liver Fibrosis Test

The Enhanced Liver Fibrosis Test is a blood test that uses an algorithm combining three biomarkers [hyaluronic acid (HA), tissue inhibitor of matrix metalloproteinase-1 (TIMP-1) and aminoterminal propeptide of procollagen type III (PIIINP)] to assess the stage and rate of progression of liver fibrosis.[57] The biomarkers are direct markers of extracellular matrix metabolism/degradation indicative of liver fibrosis. A higher concentration of the individual biomarkers leads to a higher ELF score and indicates a higher probability of more severe fibrosis.

The test requires a minimum of 0.3 ml of serum.[58] Blood samples are collected in a clinic and analysed at a central laboratory.[51]

As alcohol affects many of the variables used in the ELF test,[59,60–62] individuals should ideally be abstinent before the blood sample is taken.

The ELF test was CE marked in May 2007.[51]

### *Cautions and contraindications for the Enhanced Liver Fibrosis Test*

No cautions and contraindications have been identified by the manufacturer. However, because the test uses direct markers of fibrogenesis (HA and TIMP-1), the results will be unreliable in patients with chronic diseases characterised by fibrogenesis in organs other than the liver.[63]

## FibroTest and FibroMAX

FibroTest and FibroMAX are proprietary algorithms that use serum biochemical markers to assess the stage of liver fibrosis.[64] FibroTest (marketed in the USA as FibroSure) combines one direct marker of extracellular matrix metabolism/degradation (alpha-2-macroglobulin) and four indirect markers (apolipoprotein A1, haptoglobin, total bilirubin and gamma-glutamyl transpeptidase) with the patient's age and sex.[65]

Blood samples are not analysed at a central laboratory, but it is recommended that the individual components of the test are analysed using the same techniques and analysers as used by the reference laboratory that developed FibroTest.[66] The individual values are then entered into BioPredictive's website (www.biopredictive.com) and an algorithm calculates the results, which are presented as numeric estimates on a continuous scale ranging from 0.00 to 1.00. *Table 8* displays the correspondence proposed by Morra *et al.*[67] between FibroTest scores and the fibrosis stages identified by the METAVIR, Knodell, and Ishak fibrosis staging systems.

**TABLE 8** Relationship between FibroTest scores and fibrosis as measured by the METAVIR, Knodell, and Ishak fibrosis staging systems (after Morra R, Munteanu M, Imbert-Bismut F, Messous D, Ratziu W, Poynard T. FibroMAXTM: towards a new universal biomarker of liver disease? *Expert Rev Mol Diagn* 2007;**7**:481–90.). Reproduced with permission from Expert Reviews Ltd. All rights are reserved by Expert Reviews Ltd.

| FibroTest | METAVIR | Knodell | Ishak |
|---|---|---|---|
| 0.00–0.21 | F0 | F0 | F0 |
| 0.22–0.27 | F0–1 | F0–1 | F1 |
| 0.28–0.31 | F1 | F1 | F2 |
| 0.32–0.48 | F1–2 | F1–3 | F2–3 |
| 0.49–0.58 | F2 | F1–3 | F5 |
| 0.59–0.72 | F3 | F3 | F4 |
| 0.73–0.74 | F3–4 | F3–4 | F5 |
| 0.75–1.00 | F4 | F4 | F6 |

FibroMAX combines the components of FibroTest with alanine aminotransferase, aspartate aminotransferase, total cholesterol, triglycerides, fasting blood glucose, height and weight, and presents on one sheet the scores for FibroTest, SteatoTest (which measures hepatic steatosis) and AshTest (which measures the degree of necroinflammatory activity of alcoholic origin).[68]

FibroTest and FibroMAX are validated for use in patients with chronic viral hepatitis (B or C), ALD, and non-alcoholic fatty liver disease (NAFLD). The manufacturer recommends the use of FibroTest in patients with hepatitis B or C and FibroMAX in those with alcoholic or metabolic liver disease.[69]

Although FibroTest may be performed on blood samples from fasting or non-fasting patients, FibroMAX must be performed on samples from fasting patients.[67] As in the case of the ELF test, alcohol affects many of the variables used in FibroTest; therefore, individuals should ideally be abstinent before the blood sample is taken (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication).

FibroTest and FibroMAX are said to yield consistent results, with a reproducibility > 95%.[70,71] However, this has been questioned. A small study by Gressner *et al.*[72] suggests that FibroTest scores may vary by up to two METAVIR stages as a result of possible interlaboratory differences in measurements of the individual test components, even when the laboratories fulfil BioPredictive's technical requirements and the measurements lie within a quality-controlled, analytically acceptable range. Moreover, Poynard *et al.*'s[73] prospective analysis of discordance between the FibroTest and biopsy results in patients with hepatitis C demonstrated that critical interpretation of every FibroTest result is required in order to avoid false-positive (FP) or false-negative (FN) results.

The FibroTest and FibroMAX algorithms are patented, but not CE marked; they have not been published.[66] However, there are CE-marked kits for assessing the individual components.

### *Cautions and contraindications for FibroTest and FibroMAX*

No cautions or contraindications relating to patient safety have been identified. However, for reasons relating to test accuracy, FibroTest is not recommended for use in patients who have intercurrent illness (particularly acute inflammation, haemolysis, or Gilbert's syndrome), or who are taking medications that can cause elevated bilirubin levels[11] (these include allopurinol, anabolic steroids, some antibiotics, antimalaria medications, azathioprine, chlorpropamide, cholinergics, codeine, diuretics, adrenaline, meperidine, methotrexate, methyldopa, monoamine oxidase inhibitors, morphine, nicotinic acid, birth control pills, phenothiazines, quinidine, rifampin, steroids, sulfonamides, and theophylline[74]).

## FibroScan

FibroScan is produced by EchoSens, Paris, France, and distributed in the UK by Artemis Medical Ltd., Kent, UK[75] It is a non-invasive test that uses a specialised probe, an ultrasound and elastography system, and specialised software. The probe is placed on the skin over the right lobe of the liver and generates a mechanical pulse that sends a shear wave to the liver through the intercostal spaces.[76,77] The wave's velocity is measured by ultrasound and is determined by the stiffness of the liver, which is correlated with the degree of fibrosis.[76]

The procedure, which is painless, can be performed by trained medical or paramedical staff. Each test typically takes < 15 minutes.[75] The result is the median of 10 individual 'shots', reported as a numerical measure in kilopascal (kPa);[78] it is available immediately.[76] If a shot is unsuccessful, the machine provides no result and the whole process is deemed to have failed if no value is obtained

after ≥ 10 shots.[79] In addition, the manufacturer recommends that, to be considered reliable, successful measurements should meet the following three criteria:

- at least 10 valid shots
- a success rate (the ratio of valid shots to total number of shots) of at least 60%
- an interquartile range < 30% of the median value.[79]

FibroScan values range from 2.5 to 75 kPa; normal values are around 5.5 kPa.[80] A cut-off value of about 12.5 kPa has been proposed as optimal for discriminating between fibrosis and cirrhosis in patients with chronic hepatitis C.[81] As the threshold for cirrhosis appears to be higher in patients with ALD, it is important that disease aetiology should be established before testing.

FibroScan is claimed to measure liver stiffness in a cylinder approximately 1 cm in diameter and 4 cm long, between 25 and 65 mm below the surface of the skin.[82] The volume of this cylinder is at least 100 times that of a percutaneous liver biopsy sample, and is therefore far more representative of the whole liver, reducing the risk of sampling error. However, the claim that it may be performed satisfactorily on non-fasting patients[82] has recently been called into question by the finding that liver stiffness values increase substantially after food intake in both patients with hepatitis C infection and healthy controls.[83]

FibroScan appears to have good reproducibility. In a series of 195 patients with chronic liver disease of various aetiologies and without ascites, using the FibroScan ultrasonography guide to identify a suitable portion of the liver for examination, Fraquelli *et al.*[84] found that overall agreement between two operators was 0.98 [95% confidence interval (CI) 0.977 to 0.987], and intraobserver agreement was 0.98 for both operators. Increased BMI (> 25 kg/m$^2$), steatosis (> 24% of fatty liver cells), and histological evidence of no to mild fibrosis (METAVIR stage < F2) were all significantly associated with reduced interobserver agreement, and the most marked variability was seen in mild fibrosis (F0/F1), where interobserver agreement fell to 0.60 (95% CI 0.455 to 0.719).

FibroScan is CE marked. Because it is designed specifically to test for liver fibrosis, and does not produce an image, it cannot be used for any other diagnostic purpose.[56]

### *Cautions and contraindications for FibroScan*

No cautions or contraindications relating to patient safety have been identified. However, because elastic waves do not travel through liquids, FibroScan has no value in patients with ascites, even if it is clinically undetected.[77] Although this limitation has been claimed to be of little practical importance because the diagnosis of cirrhosis is clinically obvious in most patients with ascites,[85] patients with portal vein or hepatic vein block may present with ascites although not having appreciable hepatic fibrosis (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication).

More importantly, it is difficult or impossible to use FibroScan in obese patients because the probe is calibrated for a specific distance between the liver and the chest wall,[79] and the low-frequency vibration induced by the probe and/or the ultrasound wave can be strongly attenuated by the fatty tissue.[85] This limitation is particularly unfortunate as obese patients form an increasing proportion of the population and appear to be at increased risk of disease progression. However, a special probe with a measurement depth of 35–75 mm[17] is currently being developed for use in patients who are morbidly obese.[77] It is also impossible to use FibroScan in patients whose intercostal spaces are too narrow for the 9-mm-diameter probe to fit between the costal bones.[85]

Finally, FibroScan results may be influenced by factors such as acute liver injury, which will result in an overestimation of the degree of fibrosis.[79]

## Liver biopsy

The true gold standard for assessing the degree of liver fibrosis is histological analysis of the whole liver. As this is not possible in living patients, liver biopsy has been adopted as the reference standard.

Liver biopsy has a number of disadvantages. It is an invasive test: a hollow needle is used to remove a small sample of tissue from the patient for examination in the laboratory. It is performed in the fasting patient and is generally preceded by ultrasonography of the liver.[86] In most cases, liver biopsy is performed percutaneously, through the skin over the liver. However, in order to reduce the risk of complications associated with bleeding, it may be performed transvenously in patients with conditions such as ascites or coagulation defects, which are relatively common in ALD, and which form contraindications to percutaneous biopsy. In transvenous biopsy, a catheter carrying the biopsy needle is inserted through a vein (most commonly the jugular vein in the neck) and guided to the veins inside the liver.[87] Liver biopsy is also occasionally performed laparoscopically.[9]

Liver biopsy often involves a hospital stay. Patients should undergo liver biopsy on an outpatient basis only if they have no conditions that may increase the risk of the procedure, and then only in locations with easy access to a laboratory, blood bank, and inpatient facilities, and with staff who can observe the patient for 6 hours following the procedure. In addition, the patient should be able to return easily within 30 minutes to the hospital where the biopsy was undertaken, and should have a reliable individual to stay with on the first night post biopsy. If any of these criteria cannot be met, biopsy must be undertaken on an inpatient basis. Moreover, patients who undergo outpatient biopsy should be hospitalised if there is any significant complication, including pain that requires more than one dose of analgesic.[88] In the USA, where liver biopsy is generally performed on an outpatient basis under local anaesthetic, 1–3% of patients subsequently require hospital care for the management of complications (predominantly pain or hypotension); 60% of these complications occur within 2 hours of the biopsy and 96% within 24 hours.[86]

Because fibrosis is not evenly distributed throughout the liver, liver biopsy, which samples only 1/25,000–50,000 of the liver, carries a substantial risk of sampling error. Regev *et al.*[89] found a difference of at least one fibrosis stage between biopsy samples from the right and left hepatic lobes in 33% of patients with hepatitis C. Such sampling errors usually produce a low FP rate but a relatively high FN rate but, in the case of liver biopsy, inclusion of elements such as capsular or connective tissue will lead to overestimation of the degree of fibrosis.[16] Factors that affect the degree of sampling error include the length of the biopsy sample and the number of samples taken. Bedossa *et al.*[90] demonstrated experimentally in liver samples from patients with hepatitis C that correct staging was achieved for only 65% of cases when the biopsy sample was 15 mm in length; this figure rose to only 75% when samples 25 mm in length were analysed, and the proportion of samples which were correctly identified did not increase significantly with longer specimens. Abdi *et al.*[91] found that a single biopsy correctly identified cirrhosis in only 80% of cases (16/20) at post-mortem, but that 100% accuracy was achieved when three samples were taken,[91] while Maharaj *et al.*[92] also found that a single sample was unreliable: when three consecutive samples were taken through a single entry site, all three samples identified cirrhosis in only 50% of cases.

In addition to sampling error, liver biopsy results may be affected by intra- and inter-pathologist variation in the interpretation of samples.[93] Levels of inter- and intra-observer discrepancies as high as 10–20%[89,94] have been reported.

However, liver biopsy has the advantage that it provides information not only on liver fibrosis but also on other factors such as inflammation, necrosis, and steatosis, and permits the identification of potentially unexpected cofactors and comorbidities.[95]

### *Cautions and contraindications for liver biopsy*

As an invasive procedure, liver biopsy carries a risk of morbidity and mortality. Morbidity associated with liver biopsy may be broadly divided into minor complications, including transient discomfort at the biopsy site, post-procedural pain, and mild transient hypotension, and major complications, including more severe hypotension (systolic blood pressure < 90 mmHg), bleeding into the peritoneal or thoracic cavity, haemobilia, pneumothorax, perforation of the gall bladder or another organ, myocardial infarction, and death.[96]

Murtagh and Foerster[93] have suggested that one-third of patients who undergo liver biopsy report pain, and that complications occur in 3 per 1000 biopsies (0.3%) and death in 3 per 10,000 (0.03%). A systematic review of the adverse effects of liver biopsy undertaken as part of this project suggests that the overall rates of severe adverse events including death, and of death, in patients undergoing biopsy are higher at 0.81% (95% CI 0.71% to 0.90%) and 0.09% (95% CI 0.06% to 0.13%), respectively, for percutaneous biopsy, and 1.45% (95% CI 0.62% to 2.57%) and 0.18% (95% CI 0.02% to 0.85%), respectively, for transjugular biopsy (for details, see *Appendix 3*). The higher rates seen in patients undergoing transjugular biopsy presumably reflect the fact that they are at higher risk of adverse events because of the complications that form contraindications for percutaneous biopsy.

The systematic review found rates of any minor adverse events substantially lower than Murtagh and Foerster's[93] figure for pain alone, at 7.51% (95% CI 7.12% to 7.93%) for percutaneous biopsy and 9.14% (95% CI 6.51% to 12.37%) for transjugular biopsy. However, these results are less reliable than those for more severe adverse events because of greater variability in the definitions of minor adverse events used in the different studies (see *Appendix 3*). Moreover, the majority of included studies were retrospective audits of clinical notes that presumably only identified cases where the pain was sufficiently severe to require attention from medical or nursing staff.

The management of bleeding from internal biopsy sites has improved over the last 15–20 years because of the use of non-invasive arterial embolisation, which has a much lower associated morbidity and mortality than the open surgical procedures that were previously necessary to deal with this complication (Dr Marsha Morgan, Royal Free Hospital, London, 2010, personal communication). However, because of the particularly high risk of adverse effects in patients with alcoholic hepatitis and cirrhosis, both of which are associated with coagulation problems,[8] current UK guidance recommends that biopsy is not required for confirmation of diagnosis in patients with a high clinical suspicion of ALD in whom other causes of liver disease have been excluded using blood tests, unless it is necessary to confirm a diagnosis of AAH in order to inform specific treatment decisions.[9] In other words, in these patients, liver biopsy should be performed only if the risks it poses are outweighed by the potential benefits of improved patient outcomes consequent on changes in management which would not be possible without information that could only be obtained by biopsy.

## Hepatic venous pressure gradient measurement

Hepatic venous pressure gradient measurement is the gold standard for assessing the presence and severity of PHt in patients with alcohol-related cirrhosis, in whom splenomegaly, the clinical hallmark of PHt, is a less useful sign. HVPG measurement is an invasive procedure in which a balloon-tipped catheter is inserted into a hepatic vein via the femoral or jugular route. The pressure is measured with the balloon deflated, and again when it has been inflated to occlude

the hepatic vein. The HVPG is the difference between the two measurements; a result > 5 mmHg indicates PHt, and a result > 10–12 mmHg indicates clinically significant PHt associated with a risk of complications such as ascites, HE, and variceal bleeding.[42]

Hepatic venous pressure gradient measurement is reliable only when performed by an experienced operator.[97]

### *Cautions and contraindications for hepatic venous pressure gradient measurement*

Although invasive, HVPG measurement appears to be safe. No reports of mortality or serious complications have been published. The most common complications appear to be related to local injury to the vein used to gain access to the hepatic vein; they include leakage and haemorrhage. Other complications, such as vagal reactions and arteriovenous fistulae, are more rare. The risk of such complications is greatly reduced if deep-venous puncture is performed under Doppler ultrasound guidance, and this is particularly recommended in obese patients, or when arterial palpation is difficult. Finally, passage of the catheter through the right atrium may cause supraventricular arrhythmias; in over 90% of cases, these are self-limiting.[98]

## Upper gastrointestinal endoscopy

Upper gastrointestinal endoscopy involves the insertion of a thin, flexible viewing instrument, called an endoscope through the mouth into the oesophagus, stomach, and duodenum. It is used in patients with cirrhosis to identify medium or large varices in those areas so that prophylactic treatment may be initiated to reduce the risk of bleeding.

Endoscopy is expensive to perform.[99] Patients must have had nothing to eat or drink for 4–8 hours prior to the procedure.

### *Cautions and contraindications for upper gastrointestinal endoscopy*

Upper gastrointestinal endoscopy is invasive and unpleasant for the patient if done without deep sedation.[100] When performed for diagnostic purposes, it has a small, but not insignificant, risk of complications. These include:

- Cardiopulmonary complications related to sedation. These range from minor changes in vital signs to respiratory depression, shock/hypotension, and myocardial infarction, and account for approximately 40% of complications associated with diagnostic endoscopy.
- Infectious complications resulting either from the procedure itself or from the use of contaminated equipment; these are relatively uncommon.
- Perforation of the gastrointestinal tract; this is also relatively uncommon, but is associated with a mortality rate of approximately 25%.
- Significant bleeding; this is rare, although individuals with thrombocytopenia and/or coagulopathy are at increased risk of bleeding.[101]

### *Identification of important subgroups*

Potentially important subgroups include:

- obese patients (i.e. those with a BMI ≥ 30 kg/m$^2$ or a waist circumference > 102 cm for men and > 88 cm for women)
- patients with metabolic syndrome
- patients with concurrent alcoholic hepatitis
- patients who are not abstinent from alcohol.

All of these subgroups are important because their characteristics may affect the performance or results of the non-invasive tests assessed in this report; in addition, obese patients with ALD appear to be at increased risk of disease progression. Thus, FibroScan is more likely to fail, or to provide unreliable measurements, in patients with ALD who are obese or rotund, whereas it has been shown in apparently healthy individuals that liver stiffness measurements (LSMs) are higher in those with metabolic syndrome than in those without.[80] The inflammation associated with alcoholic hepatitis will result in an overestimation of the degree of fibrosis using FibroScan.[79] Finally, current alcohol consumption affects many components of the ELF test and FibroTest, so that the tests will provide different results in patients who are abstinent from alcohol and in those who are still drinking, even though they may have the same degree of fibrosis.

### Current usage in the NHS

None of the four tests are currently routinely performed within the NHS. However, in April 2008, FibroScan machines were said to be installed in 12 NHS hospitals (implicitly in the UK rather than in England and Wales).[75] By 2009, this figure had risen to 17 NHS hospitals and five private hospitals in England, and six NHS hospitals in Scotland; there were none in Wales.[102] The majority of FibroScan machines in place in 2008 were funded by pharmaceutical companies and/or charitable donations.[75]

In 2008, the manufacturers predicted that initial uptake would be confined to the major hepatology centres and would be limited to approximately 35 systems, but that FibroScan might subsequently expand into district general hospitals with gastroenterology departments.[75]

### Anticipated costs associated with non-invasive testing

There are no good data relating to the costs of non-invasive testing. The best data relate to the ELF test, where currently the favourable price for early adopters in the NHS is £45 per test for a volume of 100 tests a month (1200 a year), i.e. a total cost of £54,000/year (Dr Marsha Morgan, Royal Free Hospital, London 2010, personal communication). There is no indication as to what the price is likely to be for subsequent customers.

In 2007, Morra *et al.*[67] stated that the cost of FibroTest ranged from €90 to €300 and that of FibroMAX from €150 to €500. However, it seems likely that FibroTest will in the future be priced more competitively in relation to the ELF test. The cost of FibroMAX is likely to be somewhat higher than that of FibroTest as it incorporates more components and provides additional information.

The ancillary costs associated with the ELF test, FibroTest, and FibroMAX are those associated with any diagnostic blood test.

No cost data have been identified for FibroScan.

# Chapter 2

# Definition of the decision problem

## Decision problem

### *Population*

Patients with suspected liver fibrosis related to alcohol consumption.

### *Diagnostic tests under assessment*

The non-invasive tests for liver fibrosis assessed in this review take the forms of either blood tests:

- ELF test
- FibroTest
- FibroMAX

or transient elastography:

- FibroScan.

### *Reference standard tests*

The reference standard test for liver fibrosis is liver biopsy. However, HVPG measurement and upper gastrointestinal endoscopy are used to identify conditions associated with liver fibrosis, namely PHt and oesophageal varices respectively, which are of clinical importance and which can also be used as surrogates for the diagnosis of cirrhosis.

As noted in *Chapter 1, Liver biopsy*, liver biopsy carries a substantial risk of sampling error, and may also be affected by intra- and inter-observer variation in the interpretation of samples. It can, therefore, only be regarded as an imperfect reference standard.

### *Outcomes*

Relevant outcomes include:

- the diagnostic accuracy of the index test compared with the reference standard, as indicated by the area under the receiver operating characteristic curve (AUROC), for a specified fibrosis stage, and/or the sensitivity, specificity, positive predictive value, and negative predictive value, compared with the reference standard for the diagnosis of a specified fibrosis stage based on a specified cut-off point, or the data required to calculate those values [i.e. numbers of true-positive (TP), FN, true-negative (TN) and FP results]
- numbers of test failures or other withdrawals
- adverse effects associated with testing
- long-term patient outcomes (disease progression, complications related to liver disease, need for liver transplantation, mortality)
- health-related quality of life
- cost-effectiveness and cost–utility.

Studies of diagnostic or predictive accuracy are included only if they report numbers of TP, FN, TN and FP results, or measures of diagnostic accuracy (e.g. sensitivity and specificity) calculated

from those values, relating to the index test in comparison with either a reference standard test or a clinical outcome (e.g. survival or adverse clinical events).

## Discussion of outcomes measuring test accuracy

In studies of diagnostic tests, patients are generally classified by the index test (i.e. the diagnostic test being investigated) as either positive or negative (i.e. as having or not having the condition that the test is designed to identify). The same patients are also assessed using the reference standard (an established diagnostic test assumed to have 100% sensitivity and specificity), and the latter result is taken to identify the patients' true health status. If both the index test and the reference standard are positive, the result is described as a TP. If the index test is positive but the reference standard is negative, the result is termed a FP. If both the index test and the reference standard are negative, the result is termed a TN, whereas if the index test is negative and the reference standard is positive it is termed a FN. These results can be presented in a $2 \times 2$ table (*Table 9*).

The sensitivity of a test (the proportion of patients with the condition of interest who have a positive test result – i.e. TPs), indicates how good it is at correctly identifying the condition of interest, whereas its specificity (the proportion of patients without the condition of interest who have a negative test result – i.e. TNs) indicates how good it is at correctly classifying people as free of that condition. The sensitivity and specificity are expressed as simple percentages (see *Table 9*). Both FP and FN results may be harmful: FP results may cause patients to undergo further tests or receive unnecessary treatment, whereas FN results may deprive them of the treatment they need. Ideally, therefore, diagnostic tests would have both high sensitivity and high specificity. In practice, however, they often have high sensitivity at the expense of low specificity or vice versa. The optimum balance between sensitivity and specificity varies from test to test because of differences in the relative consequences of FP and FN results in the context of the condition of interest.

In practice, the situation may be more complex than indicated in *Table 9*, as the results of the index test may be neither positive nor negative, but uninterpretable. It is important that such uninterpretable results should be included in the calculations. Their inclusion will lower the sensitivity and specificity of the index test (*Table 10*). In the absence of full data relating to uninterpretable results, sensitivity analyses may be used to explore the impact of such results on sensitivity and specificity.

**TABLE 9** Calculation of sensitivity and specificity

| Index test result | Reference standard positive | Reference standard negative |
|---|---|---|
| Index test positive | TP | FP |
| Index test negative | FN | TN |
| | Sensitivity = $[TP/(TP + FN)] \times 100$ | Specificity = $[TN/(TN + FP)] \times 100$ |

**TABLE 10** Calculation of sensitivity and specificity including uninterpretable results

| Index test result | Reference standard positive | Reference standard negative |
|---|---|---|
| Index test positive | TP | FP |
| Index test uninterpretable | Uninterpretable ($U_1$) | Uninterpretable ($U_2$) |
| Index test negative | FN | TN |
| | Sensitivity = $[TP/(TP + U_1 + FN)] \times 100$ | Specificity = $[TN/(TN + U_2 + FP)] \times 100$ |

The positive predictive value of a test is the proportion of patients with positive test results who are correctly diagnosed. In clinical practice, this is generally the most important measure of test accuracy as it reflects the probability that a positive test reflects the underlying condition being tested for. Consequently, its value depends on the prevalence of the disease, which may vary, and indeed in real life situations may differ considerably from that seen in study populations. The negative predictive value is the proportion of patients with negative test results who are correctly diagnosed. The negative predictive value may be more important than the positive predictive value if the test is being used as a triage test to identify those patients who may have the condition of interest so that they may undergo further, more expensive or invasive, testing.

Although fibrosis has, for convenience, been divided into grades of severity using a number of different scales, the most common being the METAVIR scale (see *Chapter 1, Measurement of disease*), it actually forms a continuum that ranges from very mild fibrosis to severe fibrosis (cirrhosis). Consequently, the non-invasive tests for fibrosis reviewed in this report also yield continuous measurements and, therefore, diagnostic thresholds must be deliberately selected to define positive and negative results. The sensitivity and specificity of the tests will vary depending on the thresholds that have been selected. If several thresholds have been used in one data set, the diagnostic characteristics of the test in question may be illustrated using a receiver operating characteristic plot of the TP rate (sensitivity) against the FP rate (1 – specificity).[103]

The AUROC may be used as an overall measure of the performance of a surrogate test when compared with the reference standard. Bedossa and Carrat[104] state that the AUROC represents the probability that the surrogate will correctly rank two randomly chosen patients, one of whom has been classified by the reference standard as having, and the other as not having, the condition that the index and reference tests are intended to identify. Unlike sensitivity and specificity, the AUROC does not vary according to the threshold set for identification of a positive result. It therefore represents a loss of information, as compared with sensitivity and specificity, as it provides no indication of the degree to which any departure from 1.00 (a perfect result) is because of FPs, and the degree to which it is because of FNs. Moreover, Lambert *et al.*[105] have noted that, in the context of liver fibrosis, the AUROC has two main drawbacks:

- Because it assumes that the reference standard yields a binary result whereas, as has been seen, liver biopsy uses an ordinal scoring system, the fibrosis stages have to be aggregated into two groups.
- Because the proportion of each fibrosis stage in the sample can affect the AUROC, comparisons of AUROCs from populations with different distributions of fibrosis stages may be flawed. To overcome this, Poynard *et al.*[106] recommend standardising the AUROC according to the prevalence of fibrosis stages, but this method is complex and has not yet been statistically validated.[105]

As noted above, calculation of the sensitivity and specificity of the index tests involves the assumption that the reference standard has 100% sensitivity and specificity. Unfortunately, this is not true of liver biopsy (see *Chapter 1, Summary of diagnostic tests*). If the results of the reference test are not very close to the truth, the performance of the index tests will be poorly estimated.[103] In an attempt to address this problem, a number of the studies included in this review have looked in detail at cases in which the index test and reference standard have yielded discrepant results, to determine which test is more likely to have provided the correct result in each individual case.

Another serious problem relating to the use of liver biopsy as the reference standard is its relevance to the different types of non-invasive tests. On the one hand, liver biopsy is arguably a more appropriate reference standard in relation to transient elastography than in relation to tests

based on serum markers because it seeks directly to identify the degree of fibrosis present in the liver at a fixed point in time, whereas transient elastography measures the stiffness of the liver, a surrogate for fibrosis, also at a fixed point in time. By contrast, tests such as the ELF test and FibroTest, which are based on combinations of serum markers, seek to assess dynamic processes taking place within the liver. Consequently, their results may be discordant with the liver biopsy results either because the fibrotic process is highly active but fibrotic tissue has not yet developed (thus the serum tests will yield a higher result than the biopsy), or because fibrotic activity is temporarily discontinued even though there are clusters of fibrotic tissue in the liver (thus the serum tests will yield a lower result than the biopsy).[107] In such circumstances, the test results may be discordant even though both tests have yielded correct results in terms of the parameter that they set out to measure. On the other hand, it has been argued that liver biopsy is a more appropriate reference standard in relation to serum marker algorithms such as FibroTest and the ELF test, which have been designed to match histological stages of liver fibrosis as assessed by liver biopsy irrespective of biopsy accuracy, than in relation to transient elastography, which measures stiffness, a single genuine characteristic of liver tissue.[104]

### *Study design*

The best available level of evidence, with priority given to controlled studies, if available.

## Overall aims and objectives of assessment

The overall aim of the assessment is to assess the diagnostic accuracy, effect on patient outcomes, and cost-effectiveness of the specified non-invasive tests for liver fibrosis in patients suspected of having liver fibrosis related to alcohol consumption. The tests are assessed firstly as a replacement for liver biopsy, and secondly as an additional test prior to liver biopsy.

The objectives of the assessment are:

- To conduct a systematic review of the published evidence on the diagnostic accuracy and cost-effectiveness of the specified non-invasive tests for the assessment of liver fibrosis in patients suspected of having ALD.
- To develop a decision model to investigate the benefits, harms, and cost-effectiveness of non-invasive testing, either as a replacement for liver biopsy or as an additional test in the diagnostic pathway for assessing liver fibrosis. Outcomes from the model will be expressed in terms of net health benefit and cost per quality-adjusted life-year (QALY).

# Chapter 3

# Cost-effectiveness: model structure and methodology

As previously discussed, the use of non-invasive liver tests (NILTs) for assessing the fibrosis levels of patients with suspected ALD has been posited owing to the fact that the current assessment method, biopsy, is associated with morbidity and mortality. If NILTs were of sufficient accuracy in determining the level of fibrosis, then they could be used cost-effectively to either filter those patients in whom biopsy would not be appropriate, or indeed replace biopsy for some patients. Henceforth strategies aimed at filtering patients will be referred to as 'triaging strategies' and strategies aimed at replacing biopsy will be referred to as 'replacement strategies'.

The focus of the model is to evaluate the cost-effectiveness of different strategies involving NILTs when compared with biopsying all patients. Within this remit, it has been assumed that there is sufficient infrastructure for the identification and referral of patients with suspected ALD and subsequent treatment to be performed to a satisfactory level. Additional details on providing such services are contained in *Alcohol use disorders: diagnosis, assessment and management of harmful drinking and alcohol dependence*[108] and the references contained therein.

During the process of undertaking the evaluation, it became apparent that data regarding the use of NILTs within primary care were extremely scant. Pivotal studies assessing test accuracy were all undertaken in secondary or tertiary care as a gold standard (liver biopsy) was needed. As it is unethical to undertake liver biopsy in those with minimal risk of fibrosis, the trials would be subject to considerable spectrum bias and the resultant sensitivities and specificities could not be assumed to apply in primary care. Clinical experts (comprising primary, secondary and tertiary care physicians) who provided advice to the assessment groups were unanimous that there was currently insufficient evidence to appraise the tests in primary care. This advice, in conjunction with the considerable uncertainty that is prevalent with regards to the cost-effectiveness of NILTs in secondary care, was the rationale for the results of this study to focus on the cost-effectiveness of NILTs solely within secondary and tertiary care.

Owing to the uncertainty in both management and prognosis following a diagnosis of cirrhosis, advanced fibrosis or of their absence, there were few data from the systematic reviews that could be utilised within the modelling evaluation. The key parameters that were used were the sensitivity and specificity of the tests.

The population simulated within the model will be those patients that a hepatologist would wish to biopsy. Guidance from the National Institute for Health and Clinical Excellence (NICE)[9] indicates that biopsy, because of the potential for causing morbidity and mortality, should only be used when it would affect the management of the patient. It is assumed that within the model, management would only be altered where a patient had been diagnosed with cirrhosis, in which case the patient would be monitored for HCC, HE and oesophageal varices. In contrast, it is assumed that the management strategy would not change for those patients without cirrhosis, where the clinician would continue to stringently attempt to persuade the patient to become abstinent or reduce alcohol intake.

The prevention of further fibrosis (and ultimately cirrhosis) is of great importance and the model assumes that a proportion of those patients who continue to drink heavily will progress to cirrhosis, in which case the greater cost implications and reduced life expectancy will be taken into consideration.

A subset of patients will be suspected of having severe alcoholic hepatitis and/or decompensated liver disease; these will not be considered within the model. The rationale for this decision is twofold. Firstly, those patients with alcoholic hepatitis are likely to require treatment with steroids to reduce the risks of mortality; however, if the patient has decompensated cirrhosis, which can be determined by biopsy but not by any of the current NILT, then the course of steroids can cause mortality. Secondly, patients with alcoholic hepatitis will have inflammation of the liver, which can affect the validity of diagnosis provided by a NILT. There may be additional patients in whom the clinician believes that a biopsy would be unnecessary, for example where the clinical manifestations clearly indicate that the patient has cirrhosis; these patients are also not considered within the model with the assumption that the clinician would treat the patient as he or she deemed appropriate.

## The strategies analysed

Ten strategies will be considered:

1. biopsy all patients (assumed current practice)
2. triage patients with FibroScan and biopsy all those in whom cirrhosis is indicated
3. triage patients with FibroTest and biopsy all those in whom cirrhosis is indicated
4. triage patients with the ELF and biopsy all those in whom cirrhosis is indicated
5. triage patients using clinical experience and biopsy all those in whom cirrhosis is indicated
6. use FibroScan and assume that the result is definitive
7. use FibroTest and assume that the result is definitive
8. use the ELF and assume that the result is definitive
9. use clinical experience and assume that the result is definitive
10. diagnose all patients as having cirrhosis.

Strategies 5, 9 and 10 are not considered as realistic recommendations for clinical practice, but are included to provide insight regarding whether or not a formal diagnostic test is required. These strategies were of particular relevance to an earlier version of the conceptual model, in which the quality of life impacts due to continued drinking in those without cirrhosis were assumed to be greater than subsequently used in the modelling following clinical advice. Using a high decrement resulted in strategies with poor specificity being more cost-effective as the rates of abstinence are assumed to be greater in those with diagnosed cirrhosis. Although these conclusions do not apply to the final model, strategies 5, 9 and 10 were included for completeness. FibroMax was excluded from the strategies analysed as there were no data found regarding sensitivity or specificity.

The assumed current clinical practice is shown in *Figure 1*. If a patient were shown to be cirrhotic, then he or she would receive monitoring for HCC and HE and prophylactic treatment for oesophageal varices. Patients not shown to be cirrhotic would receive lifestyle advice only, which would include the strong recommendation to become abstinent or to reduce alcohol consumption. This advice would also be given to those who received monitoring.

For those patients who were not diagnosed as cirrhotic, it is assumed that a proportion will progress and become cirrhotic (incurring substantial costs and reduced life expectancy if they

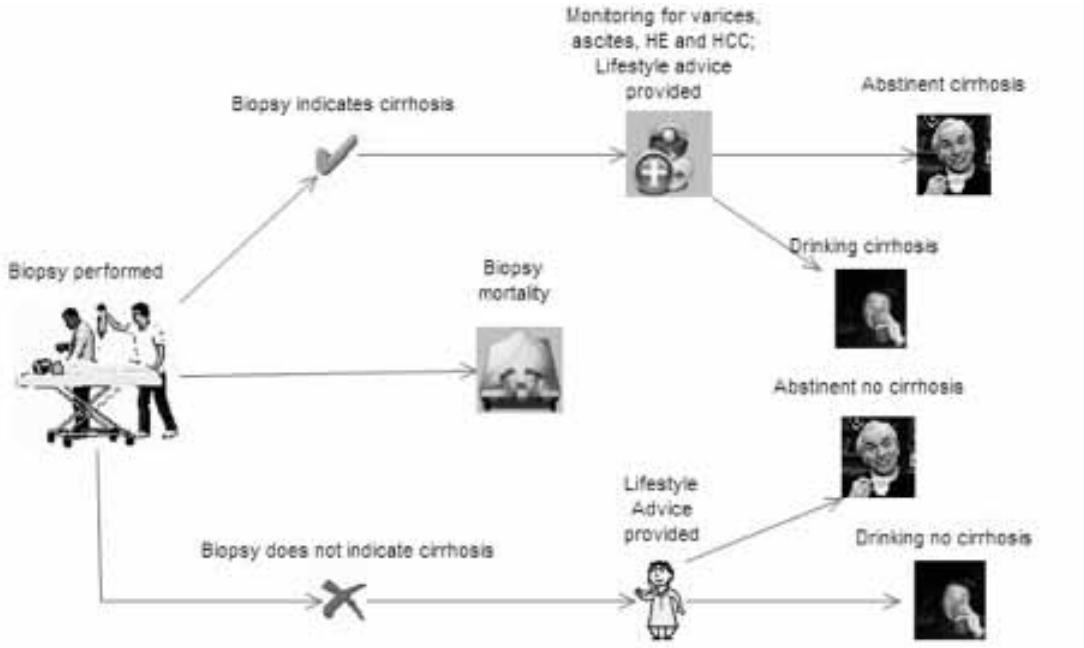



**FIGURE 1** Strategy 1: the assumed current practice.

continue to drink heavily). More detail is provided in *Chapter 5*. This pathway has not been included in *Figure 1* to maintain clarity.

For the set of strategies where the NILT would be used to triage patients, it has been assumed that the management of patients would be as shown in *Figure 2*. If the NILT indicates that there is cirrhosis, then this would be confirmed with biopsy, with those shown to be cirrhotic on biopsy receiving monitoring for HCC and HE and provided with prophylactic treatment for oesophageal varices in addition to lifestyle advice. Those patients shown to be non-cirrhotic on biopsy would receive lifestyle advice only, which would include the strong recommendation to become abstinent or to reduce alcohol consumption. This advice would also be provided to patients who are not shown to be cirrhotic by the NILT. It is assumed that the knowledge of the result of the NILT would not affect the interpretation of the biopsy result, which would provide the same diagnosis when performed immediately on a patient or following a triage strategy.

Depending on the sensitivity of the test, there is potential for patients to be diagnosed incorrectly as not having cirrhosis and not being offered appropriate monitoring for HCC, HE and prophylactic treatment for oesophageal varices.

For those patients who were not diagnosed as cirrhotic, it is assumed that a proportion will progress and become cirrhotic (incurring substantial costs and reduced life expectancy) if they continue to drink heavily. More detail is provided in *Chapter 5*. This pathway has not been included in *Figure 2* to maintain clarity.

For the analyses where it is assumed that the patient's management strategy would be determined by the NILT alone, the care pathway would be as shown in *Figure 3*.

Depending on the sensitivity of the test, there is potential for patients to be diagnosed incorrectly as not having cirrhosis and not being offered the appropriate monitoring for HCC, HE and prophylactic treatment for oesophageal varices. Additionally, depending on the specificity, there

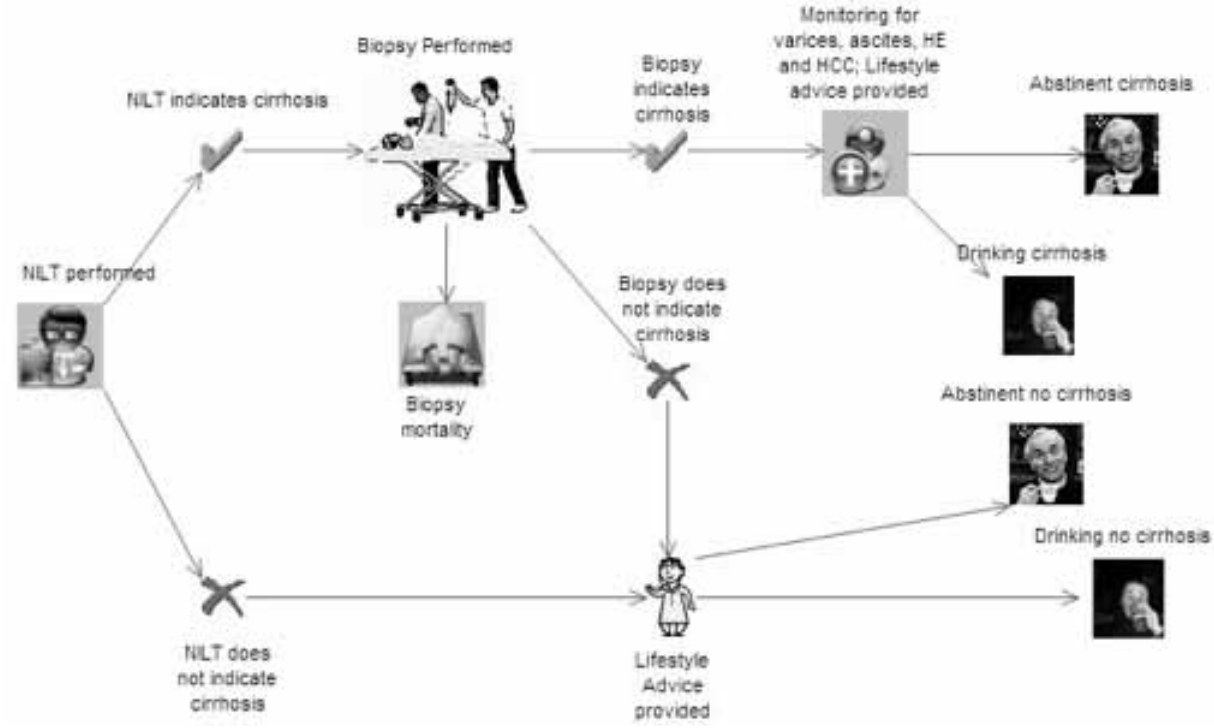



**FIGURE 2** The use of a NILT in conjunction with a biopsy.

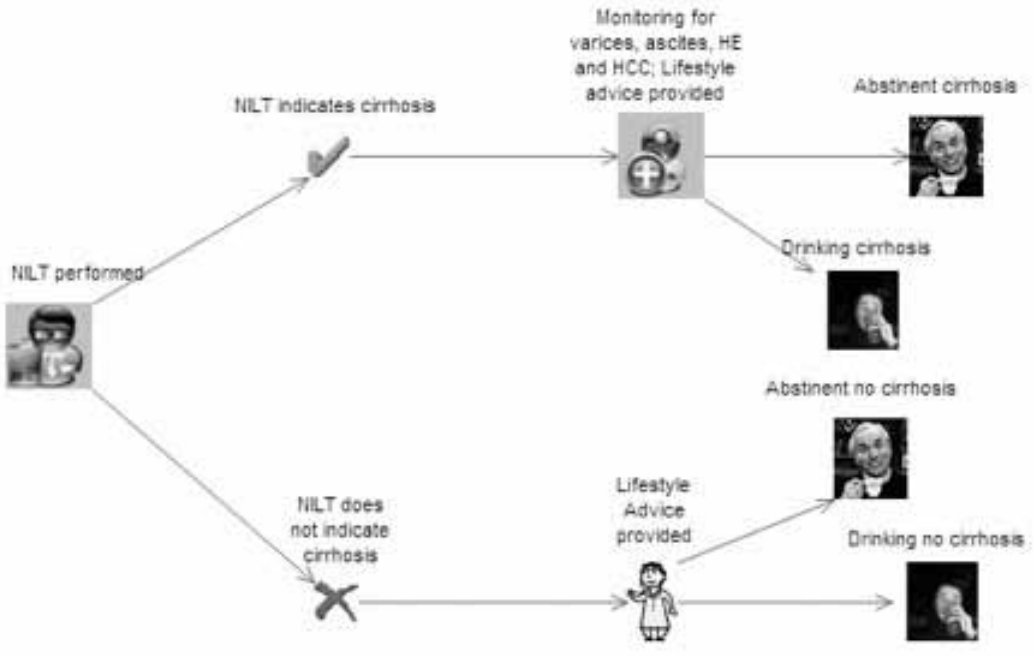



**FIGURE 3** The use of a NILT to determine patient management.

is a possibility that patients are falsely diagnosed as having cirrhosis and will have unnecessary monitoring for a considerable period of time.

For those patients who were not diagnosed as cirrhotic, it is assumed that a proportion will progress and become cirrhotic (incurring substantial costs and reduced life expectancy if they continue to drink heavily). More detail is provided in *Chapter 5*. This pathway has not been included in *Figure 3* to maintain clarity.

### *Liver transplant*

In reality, there is a possibility that patients will have a liver transplant. However, this eventuality has been omitted from the model as current evidence shows that it is of borderline cost-effectiveness. Longworth *et al.*[109] report that liver transplants in patients with ALD may not be cost-effective. However, in the recent NICE guideline[9] it was hypothesised that the cost per QALY gained estimated by Longworth *et al.*[109] may be overestimated as the selection of ALD patients for transplants may have improved, the study had not been extrapolated to patient lifetime and whether or not the full costs of pre-transplant costs should be included in the estimation of cost-effectiveness. Accordingly, the Guideline Development Group (GDG) concluded that 'that liver transplantation in its current form is likely to be cost-effective for ALD patients, when long-term benefits and modern selection practices are taken into account'.[9] As there is no conclusive evidence on whether or not liver transplantation is cost-effective, the authors of this report have assumed that the cost-effectiveness of liver transplant in an ALD population is exactly on the threshold for cost per QALY chosen by the decision makers and that whether people do, or do not, have a liver transplant will not affect the decision on whether or not the use of a NILT is cost-effective. Given the great uncertainty in the results, owing to the lack of data on key variables within the model, the authors are not uncomfortable with this assumption.

# Chapter 4

# Assessment of clinical effectiveness

## Methods for reviewing effectiveness

A systematic review was undertaken according to the general principles recommended in the quality of reporting of meta-analyses (QUOROM)[110] and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)[111] statements.

### *Identification of studies*

Extensive searches were undertaken for the comprehensive retrieval of studies of clinical effectiveness and cost-effectiveness relating to the research question. The concepts in the search strategies reflected the population and intervention categories of the Population, Intervention, Comparator, Outcome (PICO) model, namely patients with suspected liver fibrosis related to alcohol consumption and the specified non-invasive tests for the identification of fibrosis, respectively.

The search strategy comprised the following main elements:

- searching of electronic databases
- scrutiny of bibliographies of retrieved papers and previous systematic reviews
- contact with experts in the field.

#### Sources searched

The electronic databases searched included MEDLINE, EMBASE, The Cochrane Library, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Web of Knowledge (for details, see *Appendix 5*).

#### Search strategies

The MEDLINE search strategies are presented in *Appendix 5*. Search strategies for the other databases are available on request.

#### Search restrictions

Searches were not restricted by publication type, date of publication, or language.

### *Inclusion and exclusion criteria*

#### Inclusion criteria

- Participants: Patients with suspected liver fibrosis related to alcohol consumption. Studies that included patients with suspected liver fibrosis of other aetiologies were included if data relating to patients with suspected alcohol-related disease could be extracted separately.
- Intervention: One of the specified non-invasive tests for liver fibrosis, namely:
  - the ELF test
  - FibroTest
  - FibroMAX
  - FibroScan.
- Comparators: The primary comparator, or reference standard, was liver biopsy for the identification of liver fibrosis. Secondary reference standards were tests used to identify

conditions associated with liver fibrosis, namely HVPG measurement for PHt, and upper gastrointestinal endoscopy for the identification of oesophageal varices.

- Outcome measures: The primary outcome measure was the diagnostic accuracy of the index test compared with the reference standard in distinguishing patients with significant fibrosis, defined as METAVIR stages F2–F4, from patients without significant fibrosis, defined as METAVIR stages F0–F1. Other outcome measures were:
  - the diagnostic accuracy of the index test compared with the reference standard in distinguishing patients with cirrhosis (METAVIR stage F4) from patients without cirrhosis (METAVIR stages F0–F3)
  - the diagnostic accuracy of the index test compared with the reference standard in distinguishing patients with moderate-to-severe fibrosis (METAVIR stages F3–F4) from patients without moderate-to-severe fibrosis (METAVIR stages F0–F2)
  - the diagnostic accuracy of the index test compared with the reference standard in distinguishing patients with fibrosis (METAVIR stages F1–F4) from patients without fibrosis (METAVIR stage F0)
  - the diagnostic accuracy of the index test compared with the reference standards in distinguishing patients with and without the complications of fibrosis (PHt and oesophageal varices)
  - the number of patients requiring referral to secondary care
  - the number of patients requiring liver biopsy
  - the number of patients giving up alcohol, or significantly reducing alcohol consumption, following receipt of a test result
  - long-term patient outcomes (disease progression, complications related to liver disease, need for liver transplantation, mortality)
  - adverse effects of testing
  - health-related quality of life.

Only studies of the index tests that reported data relating to one of the outcome measures in relation to the population of interest were included in the review of clinical effectiveness. However, this criterion was relaxed for consideration of adverse events, where wider searches were undertaken to allow the inclusion of data relating to studies of the adverse effects of diagnostic venepuncture or transient elastography (see *Appendix 5*).

- Study design: the best available level of evidence, with priority given to controlled studies, if available.

### Exclusion criteria

The following publication types were excluded from the review:

- animal models
- preclinical and biological studies
- narrative reviews, editorials, and opinions.

Systematic reviews of primary studies were excluded from the review of clinical effectiveness, but were scanned for potential additional relevant studies.

In addition, studies were excluded if:

- they were considered methodologically unsound (specifically, if the reference standard was used in only a subset of study participants and the selection criteria used to identify that subset were not clear)

- they were published as meeting abstracts only, and insufficient methodological details were reported to allow critical appraisal of study quality
- they were meeting abstracts that had been superseded by later publications and did not contain any additional data.

#### Sifting

The references identified by the literature searches were sifted in three stages by a single reviewer. They were screened for relevance first by title and then by abstract. Those papers that seemed, from their abstracts, to be relevant were then read in full, as were all potentially relevant papers for which abstracts were not available. At each step, studies that did not satisfy the inclusion criteria were excluded.

### *Data extraction strategy*

Data were extracted by one reviewer to customised data extraction forms. Where multiple publications of the same study were identified, data were extracted and reported as a single study.

### *Critical appraisal strategy*

Study quality was assessed using a modified version of the quality assessment of diagnostic accuracy studies (QUADAS) checklist,[112] a validated tool designed to assess the internal and external validity of studies of diagnostic accuracy. Definitions of some scoring criteria were adapted from the systematic review by Friedrich-Rust *et al.*[113] (for details, see *Appendix 6*). Where a study was reported in more than one publication, its quality was assessed on the basis of the combined data from all relevant publications.

The quality assessment of studies included in the review of clinical effectiveness was carried out by one researcher. Blinding of the quality assessor to author, institution, or journal was not considered necessary.[114,115]

### *Methods of data synthesis*

Studies that met the review's entry criteria were eligible for inclusion in meta-analyses if this was appropriate in terms of comparability of the study populations, outcomes, and diagnostic thresholds, if the studies were unlikely to be biased,[116] and if the numbers of TP, FP, TN, and FN results for each study were reported or could be obtained from the study authors. However, because of the degree of heterogeneity and the unavailability of full data from some studies, meta-analysis was not in fact considered appropriate. The presentation of results is therefore limited to a narrative review.

Where they were not reported by the original investigators, if data were available for the numbers of TPs, TNs, FPs and FNs, the reviewers independently calculated sensitivity, specificity and positive and negative predictive values, with CIs. This was done using beta distributions (alpha = TPs and beta = FNs for specificity; alpha = TNs and beta = FPs for specificity). If any number was < 5, a non-informative prior of 0.5 (equivalent to Jeffreys' prior) was added to both the alpha and beta parameter.

## Results

### *Quantity and quality of research available*

The electronic literature searches identified 4039 potentially relevant citations. Of these, 3829 were excluded at the title or abstract stage, leaving 210 that were obtained for examination of the full text, together with five additional relevant articles that had been identified from other sources. Two of these five articles, those by Janssens *et al.*[117] and Mueller *et al.*,[118] had been

published too recently to be identified by the electronic searches; they superseded abstracts[119,120] that had been identified by those searches. A further two articles, by Rosenberg *et al.*[121] and Melin *et al.*,[122] were not identified by the electronic searches because they were not appropriately indexed in the electronic databases; they were supplied by the relevant manufacturers together with an unpublished paper by Parkes *et al.*[57]

One hundred and ninety-eight citations were excluded at the full-text stage, leaving 17 articles that were included in the review (*Figure 4*). These 17 articles related to 14 studies: one study of the ELF test,[57,121] four studies of FibroTest,[13,123–126] eight studies of FibroScan[97,117,118,122,127–131] and one study which used both FibroTest and FibroScan.[132] No studies of FibroMAX were identified.

## Number and type of studies included

The majority of articles that met the review inclusion criteria reported cross-sectional studies intended to confirm the performance of one of the non-invasive tests against a reference standard (liver biopsy, HVPG measurement, or endoscopic identification of oesophageal varices). A minority had a cohort design, following patients over time to assess how well the non-invasive test predicted adverse clinical outcomes. There were no randomised controlled trials (RCTs).

## Number and type of studies excluded, with reasons

As may be seen from *Quantity and quality of research available*, a substantial number of the citations identified by the electronic searches related to studies that were excluded as part of the sifting process because they did not meet the inclusion criteria. Details are therefore given only of those citations that were excluded after a full reading, and then only if they were excluded for a reason other than a simple failure to meet the inclusion and exclusion criteria. Such citations are listed in *Appendix 7,* together with the reasons for their exclusion.

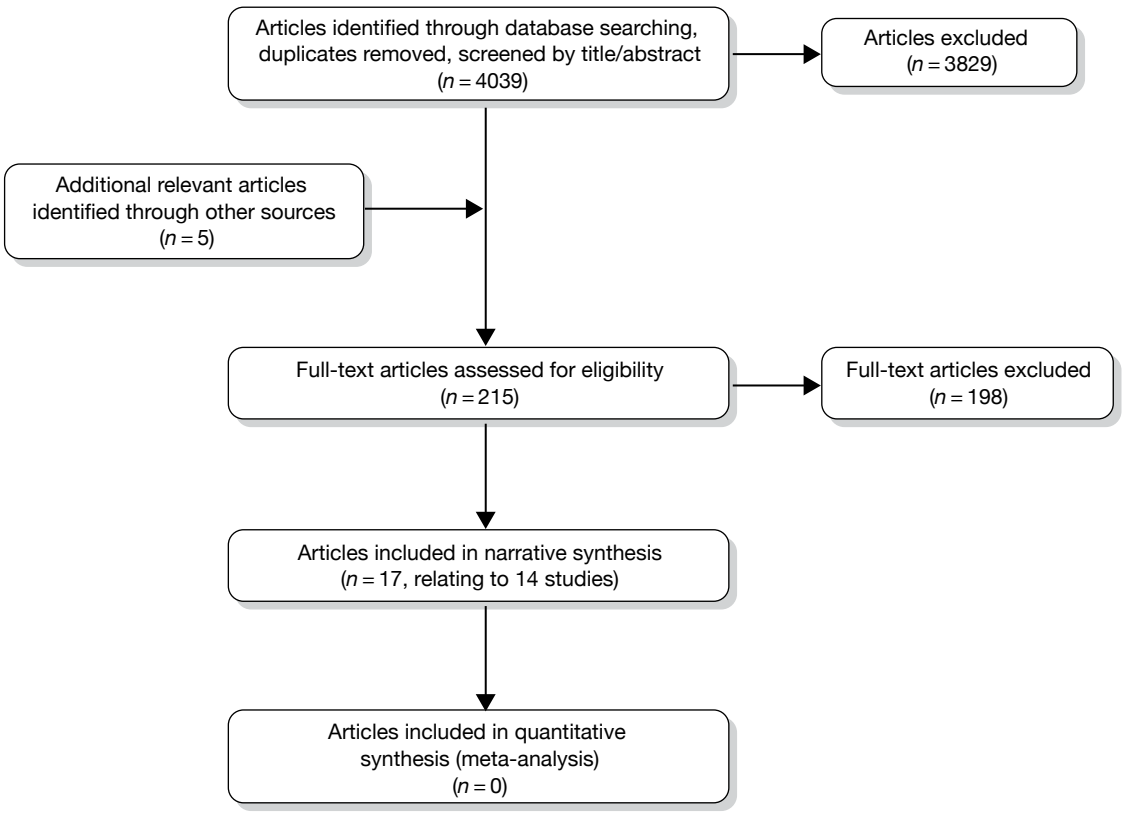



**FIGURE 4** Clinical effectiveness: summary of study selection and exclusion.

## Study characteristics

### *Enhanced Liver Fibrosis Test*

No studies were identified that assessed the ELF test as such. However, one study[57] evaluated the European Liver Fibrosis Test, which was said to be essentially identical to the ELF test except for the inclusion of age in the algorithm. This study was, therefore, deemed to meet the inclusion criteria. Data from this study have been published in two articles.[57,121] The first article, by Rosenberg *et al.*,[121] assessed diagnostic test accuracy compared with liver biopsy in patients with chronic liver disease, some of whom had ALD. The second article, by Parkes *et al.*,[57] evaluated the ability of the test to predict survival and relevant adverse events in the cohort of patients with chronic liver disease of various aetiologies enrolled in the original study in English hepatology centres. The primary outcome measure used by Parkes *et al.*[57] was the first post-recruitment liver-related clinical event, defined as liver-related death, ascites, encephalopathy, oesophageal variceal haemorrhage confirmed by endoscopy, liver transplantation, or HCC. The presence of varices without haemorrhage was not included as an outcome because of the possibility that differences in the practice of endoscopy in the different centres may lead to ascertainment bias. *Table 11* provides further details of study design.

At first sight, the data relating to the number of patients with ALD included in the analyses by Rosenberg *et al.*[121] and Parkes *et al.*[57] are confusing. One thousand and twenty-one patients with chronic liver disease of any aetiology were eligible for inclusion in the original study, and 921 were recruited. Of these, 621 formed the training or derivation cohort used to identify the optimum combination of markers and algorithm (the European Liver Test) which was then assessed in the remaining 300 patients (the validation cohort). All 64 patients with ALD were included in the validation cohort (Professor William Rosenberg, University College London, 2010, personal communication). The follow-up study by Parkes *et al.*[57] states that 85 of the patients enrolled in the study in English centres alone had ALD (i.e. more patients with ALD than were said by Rosenberg *et al.* to have been included in the original study); this apparent discrepancy is attributed to the fact that some patients originally believed to have liver disease of a different aetiology were later found to have ALD (Professor William Rosenberg, University College London, 2010, personal communication).

### *FibroTest*

Two studies of FibroTest were identified that specifically recruited patients with known or suspected ALD. Nguyen-Khac *et al.*[132] evaluated diagnostic test accuracy compared with liver biopsy, but did not include the full spectrum of ALD: patients with known or decompensated cirrhosis were excluded on the basis that they did not require further investigation. Naveau *et al.*[123] compared FibroTest and liver biopsy results in patients hospitalised either for complications of cirrhosis or for alcoholism. The study also assessed the ability of FibroTest to predict 5- and 10-year survival in 218 of the 292 patients (75%) enrolled in the study of test accuracy and followed up for a median period of 8.2 years (range 5 days to 11.8 years).[124]

A further three studies,[13,125,126] all by Thabut *et al.*, evaluated FibroTest in patients with liver disease of mixed aetiology, including ALD. To avoid any risk of double-counting, clarification was obtained from the author that no patient was included in more than one of these studies (Dr Dominique Thabut, Hôpital Pitié-Salpêtriére, Paris, 2010, personal communication). One study[13] assessed the ability of FibroTest to identify PHt in patients undergoing transjugular liver biopsy for clinical reasons and also compared FibroTest and liver biopsy results in these patients. A second study[125] assessed its ability to predict the presence of oesophageal varices in patients with chronic liver disease. The third study[126] assessed FibroTest's predictive value in relation to survival at 2 months and 6 months in patients with severe cirrhosis.

*Table 12* provides further details of study design.

**TABLE 11** Characteristics of included analyses of the ELF test study

| Study | Country/type of centre | Recruitment dates | Reference standard (fibrosis staging system) | Minimum biopsy length/minimum number of portal tracts/maximum time between index and reference test | Study design | Total patients with ALD/total study population | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|
| Rosenberg 2004[121] | Multinational/NR | 1998–2000 | Liver biopsy (Scheuer and Ishak systems, modified for conditions other than chronic viral or immune hepatitis) | 12 mm/≥ 5 portal tracts/tests performed on same day | Cross-sectional/ prospective, consecutive | Eligible: 64/1021<br>Assessed: 64[a]/300 | Patients aged 18–75 years due to undergo liver biopsy to investigate chronic liver disease (defined as abnormal liver function tests for > 6 months) | Disorders associated with extrahepatic fibrosis (including rheumatic, renal, or lung disease); CVD or cancer; advanced cirrhosis with evidence of decompensation (Child–Pugh class C); hepatocellular carcinoma; drug-induced liver disease; regular aspirin consumption; inability to provide informed consent |
| Parkes 2010[57] | England/seven hepatology centres in secondary or tertiary care | 1998–2000 | Clinical outcomes | Not applicable | Cohort/ prospective, consecutive | Eligible: NR/NR<br>Assessed: 85/457 | Patients aged 18–75 years due to undergo liver biopsy to investigate chronic liver disease (defined as abnormal liver function tests for > 6 months) | Disorders associated with extrahepatic fibrosis (including rheumatic, renal, or lung disease); CVD or cancer; advanced cirrhosis with evidence of decompensation (Child–Pugh class C); hepatocellular carcinoma; drug-induced liver disease; regular aspirin consumption; inability to provide informed consent |

CVD, cardiovascular disease; NR, not reported.

a Professor William Rosenberg, University College London, 2010, personal communication.

**TABLE 12** Characteristics of included analyses of FibroTest studies

| Study | Country | Recruitment dates | Reference standard (fibrosis staging system) | Minimum biopsy length/minimum number of portal tracts/maximum time between index and reference test | Study design | Total patients with ALD/total study population | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|
| Naveau 2005[123] | France | February 1996 to May 2000 | Liver biopsy (modified METAVIR) | Minimum length and number of portal tracts NR [mean length 15 (SE 0.5) mm; mean number of portal tracts 14.4 (SE 0.7)]/31 days | Cross-sectional/ prospective, consecutive | Eligible: 292/292<br>Assessed: 221/221 | Hospitalised for complications of cirrhosis (jaundice 12%, ascites 9%, digestive tract haemorrhage 3%) or alcoholism without complications of cirrhosis (76%); self-reported daily alcohol consumption in previous year ≥ 50 g pure ethanol (mean 146 ± 80 g/day for 17 years); available serum levels; 'consistent' liver biopsy | Concomitant liver diseases including hepatitis B or C; HIV antibodies; immunosuppression; serum or biopsy unavailable; > 31 days between biopsy and collection of serum |
| Naveau 2009[124] | France | As above | 5- and 10-year survival | Not applicable | Cohort/ prospective, consecutive | Eligible: 292/292<br>Assessed: 218/218 | | |
| Nguyen-Khac 2008[132] | France | April 2005 to January 2007 | Liver biopsy (METAVIR) | Minimum length and number of portal tracts NR (mean length 12.2 ± 3 mm; mean number of portal tracts 7.8 ± 2.7)/tests performed on same day | Cross-sectional/ prospective | Eligible: 160/160<br>Assessed: 103/103 | Age ≥ 18 years; alcohol intake > 50 g/day for > 5 years; consulting hospital department of hepato-gastroenterology, alcoholism or internal medicine for detoxification and/or inpatient rehabilitation | Chronic viral hepatitis B or C; severe AAH (Maddrey score ≥ 32); known or decompensated alcoholic cirrhosis (ascites, oesophageal varices, prothrombin time < 70%, or imaging evidence of cirrhosis); pregnancy; refusal to undergo liver biopsy; failure to obtain valid transient elastography measurements; absence of consent |

*continued*

**TABLE 12** Characteristics of included analyses of FibroTest studies *(continued)*

| Study | Country | Recruitment dates | Reference standard (fibrosis staging system) | Minimum biopsy length/minimum number of portal tracts/maximum time between index and reference test | Study design | Total patients with ALD/total study population | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|
| Thabut 2003[125] | France | NR | Endoscopy | Not applicable/<6 months | Cross-sectional/ consecutive | Eligible: 57/120[a]<br>Assessed: 57/120 | Chronic liver disease | History of liver transplantation |
| Thabut 2007a[13] | France | Unspecified 8-month period | Liver biopsy (METAVIR), HVPG measurement | Minimum length and number of portal tracts NR/tests performed on same day | Cross-sectional/ prospective, consecutive | Eligible: NR/147<br>Assessed: 66/130 | Undergoing transjugular liver biopsy for clinical reasons | History of liver transplantation; cardiac hepatopathy; pre-sinusoidal and extrahepatic PHt; renal disease |
| Thabut 2006[126] | France | NR | Survival at 2 months and 6 months | Not applicable | Cross-sectional/ prospective | Eligible: NR/NR<br>Assessed: 175/224[a] | Severe cirrhosis; enrolled in placebo-controlled trial of pentoxifylline (Trental®, Sanofi-Aventis) | NR |

NR, not reported.
a Dr Thierry Poynard, Hôpital Pitié-Salpêtriére 2010, personal communication.

### FibroMAX

No relevant studies of FibroMAX were identified.

### FibroScan

Six studies[117,118,122,127,128,132] were identified that specifically recruited patients with known or suspected ALD and assessed the diagnostic test accuracy of FibroScan relative to liver biopsy in these patients. In the studies by Kim *et al.*,[127] Mueller *et al.*,[118] Nahon *et al.*[128] and Nguyen-Khac *et al.*,[132] all patients who underwent FibroScan were also biopsied. However, in the studies by Janssens *et al.*[117] and Melin *et al.*,[122] biopsy was only undertaken in a subset of patients who underwent FibroScan. Janssens *et al.*[117] used FibroScan to identify those patients requiring alcohol detoxification or rehabilitation who had a score of ≥ 9.5 kPa; this threshold was chosen as it was thought to be indicative of severe fibrosis (F3–F4). Data relating to the test accuracy of FibroScan compared with biopsy and HVPG were therefore available only for patients with a FibroScan score of ≥ 9.5 kPa who then consented to liver biopsy and in whom both tests were conducted successfully. Similarly, Melin *et al.*[122] sought to compare the accuracy of FibroScan with biopsy in patients being treated for alcohol withdrawal who had a FibroScan score higher than 13 kPa; this threshold was apparently chosen because it was considered to be the appropriate threshold for the diagnosis of cirrhosis in patients with hepatitis C. Only 41 patients met this criterion; three of these refused biopsy and a further three had contraindications to biopsy.

A further three studies[97,129,130] evaluated the diagnostic test accuracy of FibroScan in patients with liver disease of mixed aetiology, including ALD. Bureau *et al.*[129] studied the ability of FibroScan to predict significant PHt in patients undergoing transjugular liver biopsy for clinical reasons. Lemoine *et al.*[97] and Nguyen-Khac *et al.*[130] both specifically recruited patients with cirrhosis. Lemoine *et al.*[97] assessed the ability of FibroScan relative to HVPG measurement to predict significant PHt in patients with compensated cirrhosis, whereas Nguyen-Khac *et al.*[130] assessed its ability relative to upper intestinal endoscopy to predict the presence of large oesophageal varices in patients with cirrhosis of unspecified severity.

*Table 13* provides further details of study design.

## Study quality

*Figures 5–9* provide an overview of the methodological quality of the included studies.

As may be seen, few studies presented results relating to a representative spectrum of patients suspected of having ALD. The majority recruited patients with relatively severe disease. Bureau *et al.*,[129] Rosenberg *et al.*[121] and Thabut *et al.* 2007a[13] recruited patients who were due to undergo liver biopsy for clinical reasons, whereas Lemoine *et al.*,[97] Nguyen-Khac *et al.*[130] and Thabut *et al.* 2007b[126] recruited patients known to have cirrhosis, Naveau *et al.*[123] recruited patients hospitalised for complications of cirrhosis or alcoholism, and Mueller *et al.*[118] and Thabut *et al.* 2003[126] recruited those known to have, rather than suspected of having, chronic liver disease. Two studies, those by Janssens *et al.*[117] and Melin *et al.*,[122] recruited more representative populations, but displayed partial verification bias, using the reference standard only in patients scoring above a specific threshold on the index test (FibroScan in both cases).

In studies of test accuracy, it is clearly important that the interval between the performance of the index and reference tests should be as short as possible, to minimise the possibility of the patient's condition altering significantly between the tests. However, several studies allowed a delay of > 2 weeks between the index test and reference standard. Naveau *et al.*[123] allowed an interval of up to 1 month, whereas Kim *et al.*[127] allowed the interval between transient elastography and liver biopsy to be as much as 92 days, and Thabut *et al.* 2003[125] included patients in whom endoscopy

**TABLE 13** Characteristics of included analyses of FibroScan studies

| Study | Country | Recruitment dates | Details of FibroScan assessment | Reference standard (fibrosis staging system) | Minimum biopsy length/minimum number of portal tracts/maximum time between index and reference test | Study design | Total patients with ALD/total study population | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|---|
| Bureau 2008[129] | France | 15 November 2005 to 15 October 2006 | Right lobe of liver assessed through intercostal spaces. Reported result median of at least 10 validated measurements with a success rate ≥ 60% | Liver biopsy (METAVIR[133]); HVPG measurement; upper gastrointestinal endoscopy (patients with cirrhosis only) | Minimum length and number of portal tracts NR [mean length 15.5 (SE 7.1) mm]/tests performed on same day | Cross-sectional/ prospective, consecutive | Eligible: 51/150<br>Assessed: NR/144 | Undergoing transjugular liver biopsy because of chronic liver abnormalities | Portal-vein thrombosis; Budd–Chiari syndrome; antiviral therapy or pharmacological treatment known to modify portal pressure |
| Janssens 2010[117] | Belgium | 1 January 2006 to 29 February 2008 | Reported result median of at least 10 validated measurements, with a success rate ≥ 60% and interquartile range < 30%. Procedure said to be performed by an experienced examiner | Transjugular liver biopsy (METAVIR[133]); HVPG measurement | 15 mm; ≥ 6 portal tracts/circa 2 weeks | Cross-sectional/ apparently prospective | Eligible: 255/255<br>Biopsy offered to patients with FS score ≥ 9.5 kPa (72/239), and performed in 49 | Admitted to unit for alcohol detoxification and rehabilitation; self-reported minimum daily alcohol intake ≥ 70 g/day | Declined to be rehospitalised for second week of programme; declined FS; FS unsuccessful; refused liver biopsy after successful FS |
| Kim 2009[127] | Republic of Korea | August 2006 to April 2007 | Right lobe of liver assessed following ultrasound to identify appropriate location. Reported result apparently median of at least 10 validated measurements with a success rate ≥ 30% | Liver biopsy (Batts–Ludwig[134]) | 10 mm; > 10 periportal tracts/92 days | Cross-sectional/ probably prospective and consecutive | Eligible: 51/51<br>Assessed: 45/45 | ALD (alcohol consumption ≥ 80 g/day (mean 150.7 ± 60.4 g, range 80–320 g; duration of alcohol consumption 20.8 ± 6.8 years, range 8–40 years) | Ultrasound evidence of ascites; factors other than viral liver diseases or ALD that could cause chronic liver disease; AAH |

| Study | Country | Recruitment dates | Details of FibroScan assessment | Reference standard (fibrosis staging system) | Minimum biopsy length/minimum number of portal tracts/maximum time between index and reference test | Study design | Total patients with ALD/total study population | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|---|
| Lemoine 2008[97] | France | January 2004 to September 2006 | Right lobe of liver assessed through intercostal spaces. Reported result median of at least 10 validated measurements, with success rate ≥ 70% and interquartile range ≤ 30% of the median value. Measurements said to be obtained by two experienced operators | HVPG measurement<br>Liver biopsy (Chevallier *et al.*[135]) | Minimum length and number of portal tracts NR (mean length 15 mm)/tests performed on same day | Cross-sectional/ prospective, consecutive | Eligible: 48/92<br>Assessed: apparently 48/92 | Histologically proven compensated cirrhosis (Child–Pugh class A) related to HCV or ALD (the latter defined by alcohol intake ≥ 50 g/day for ≥ 10 years) undergoing transjugular liver biopsy with HVPG measurement and transient elastography on the same day | Liver disease of other aetiologies, in particular hepatitis B or mixed aetiology (alcohol and HCV, or HIV co-infection); portal-vein thrombosis; ongoing beta-blocker therapy; ongoing or recent (< 6 months) antiviral therapy; recent gastrointestinal bleeding or endotherapy; liver cancer; cardiac failure |
| Melin 2005[122] | France | Unspecified 6-month period | Right lobe of liver assessed through intercostal spaces following ultrasound to identify appropriate location. Reported result median of up to 10 validated measurements | Liver biopsy | NR | Cross-sectional/ prospective, consecutive | Eligible: 245/245<br>Biopsy offered to patients with FS score ≥ 13 kPa (41/227) and performed in 35 | Patients seeking treatment for alcohol withdrawal; FS result ≥ 13 kPa | NR |
| Mueller 2010[118] | Germany | NR | Right lobe of liver assessed through intercostal spaces. Reported result median of 10 procedures with a success rate ≥ 60% | Liver biopsy (Kleiner[136]) | 15 mm/minimum number of portal tracts and maximum time between tests NR | Cross-sectional (validation cohort only)/ prospective, consecutive | Eligible: 106/106<br>Assessed: 101/101 | Histologically staged ALD (alcohol intake > 60 g/day[120]) | Biopsy < 15 mm; ultrasound results indicative of extrahepatic cholestasis, liver congestion, or liver tumours |

*continued*

**TABLE 13** Characteristics of included analyses of FibroScan studies *(continued)*

| Study | Country | Recruitment dates | Details of FibroScan assessment | Reference standard (fibrosis staging system) | Minimum biopsy length/minimum number of portal tracts/maximum time between index and reference test | Study design | Total patients with ALD/total study population | Inclusion criteria | Exclusion criteria |
|---|---|---|---|---|---|---|---|---|---|
| Nahon 2008,[128] Nahon 2007[131] | France | November 2005 to November 2006 | Reported result median of at least 10 validated measurements with a success rate ≥ 50%. Procedure said to be performed by an experienced examiner | Liver biopsy (Brunt *et al.*[137] Chevallier *et al.*[135]) | 10 mm/minimum number of portal tracts NR/tests performed on same day | Cross-sectional/ prospective, consecutive | Eligible: 174/174<br>Assessed: 147/147 | New patients referred to hepatology units with suspected ALD (alcohol intake > 80 g/day for > 10 years) | Ascites; HIV; hepatitis B or C; refused FS or liver biopsy; FS and liver biopsy not performed on same day |
| Nguyen-Khac 2008[132] | France | April 2005 to January 2007 | Reported result median of at least 10 validated measurements with a success rate ≥ 60% | Liver biopsy (METAVIR[133]) | Minimum length and number of portal tracts NR (mean length 12.2 ± 3 mm; mean number of portal tracts 7.8 ± 2.7)/tests performed on same day | Cross-sectional/ prospective | Eligible: 160/160<br>Assessed: 103/103 | Age ≥ 18 years; alcohol intake > 50 g/ day for > 5 years; consulting hospital department of hepato-gastroenterology, alcoholism or internal medicine for detoxification and/or inpatient rehabilitation | Chronic viral hepatitis B or C; severe AAH (Maddrey score ≥ 32); known or decompensated alcoholic cirrhosis (ascites, oesophageal varices, prothrombin time < 70%, or imaging evidence of cirrhosis); pregnancy; refusal to undergo liver biopsy; failure to obtain valid transient elastography measurements; absence of consent |
| Nguyen-Khac 2009[130] | France | NR | NR | Upper intestinal endoscopy | Not applicable/ maximum time between tests NR | Not clear | Eligible: 103/183<br>Assessed: apparently 103/183 | Cirrhosis | NR |

FS, FibroScan; NR, not reported.

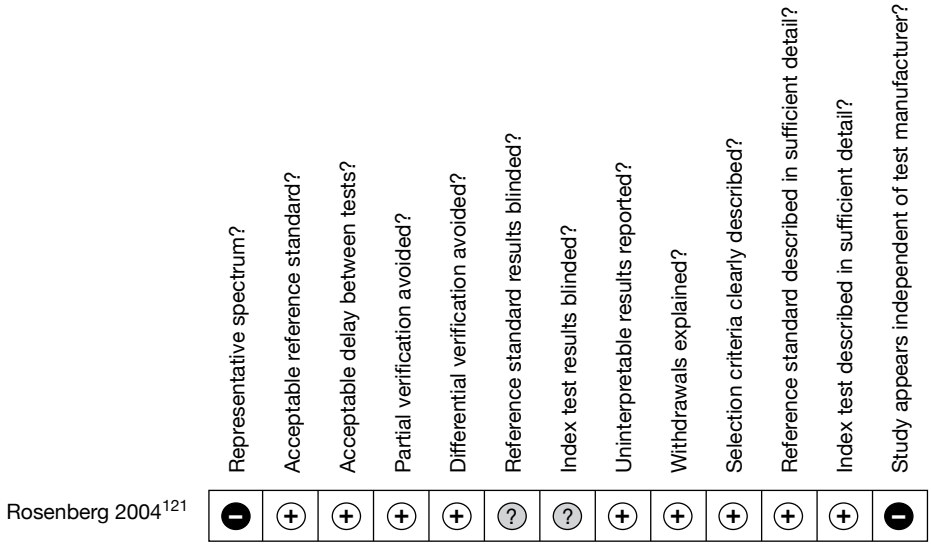



**FIGURE 5** The ELF test: methodological quality summary. Review authors' judgements about each methodological quality item. –, no; +, yes; ?, unclear.

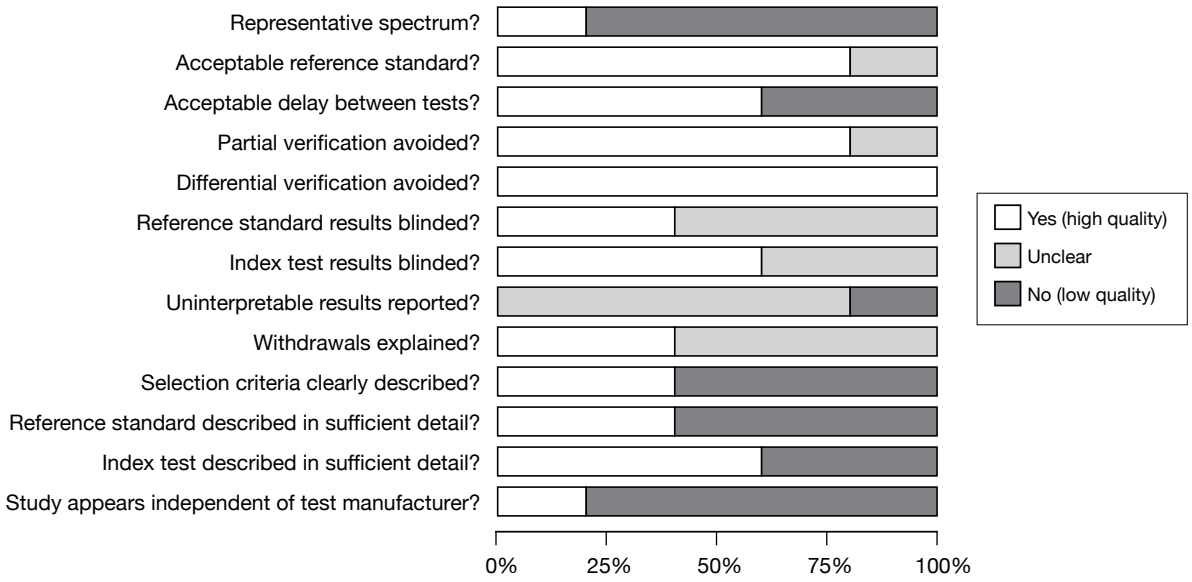



**FIGURE 6** FibroTest: methodological quality graph. Review authors' judgements about each methodological quality item presented as percentages across all included studies.

was performed up to 6 months before or after FibroTest, although in this case the mean interval was only 5 days.

In approximately half of the included studies, it was not clear whether the reference standard results were interpreted without knowledge of the results of the index test ('reference standard results blinded') and vice versa ('index test results blinded'). The remaining studies stated that blinding was used for either one or both tests. The index test was usually well described, but in many cases the execution of the reference standard was not described in sufficient detail to permit replication of the test precisely as performed by the study investigators.

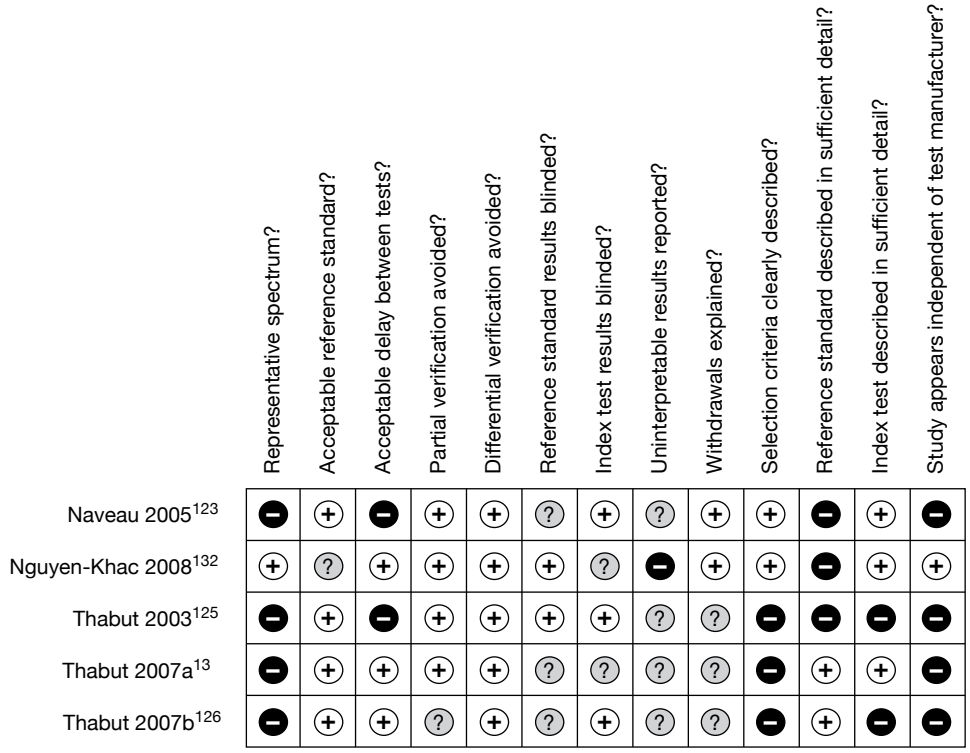



**FIGURE 7** FibroTest: methodological quality summary. Review authors' judgements about each methodological quality item for each included study. –, no; +, yes; ?, unclear.

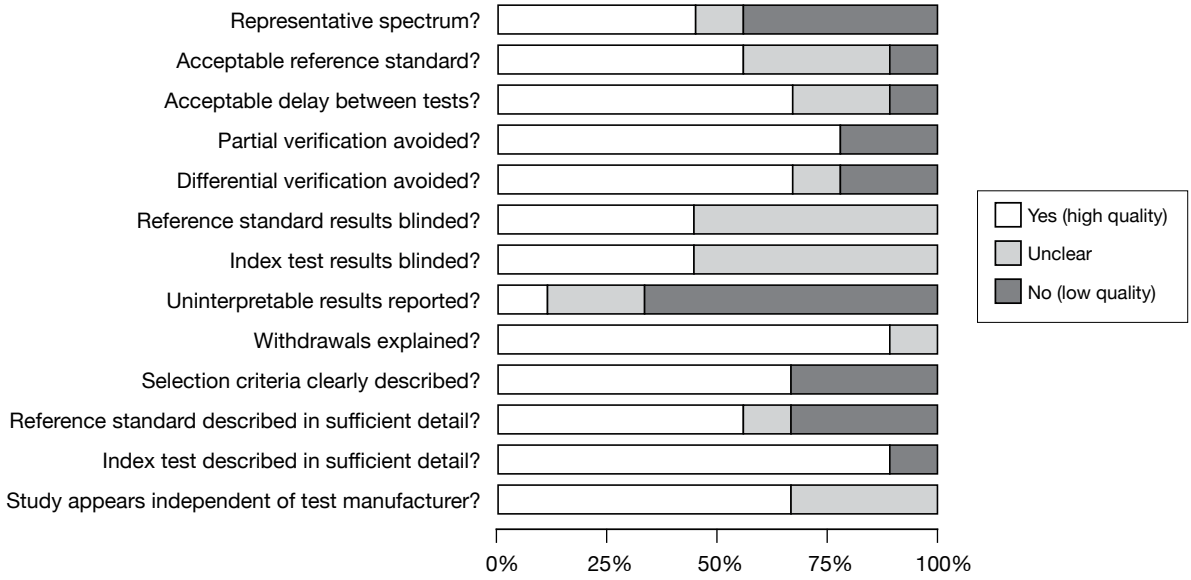



**FIGURE 8** FibroScan: methodological quality graph. Review authors' judgements about each methodological quality item presented as percentages across all included studies.

Uninterpretable/intermediate results were generally poorly reported. Two studies, Rosenberg *et al.*'s[121] study of the ELF test and Lemoine *et al.*'s[97] FibroScan study, stated that patients were recruited prospectively and consecutively, and reported no uninterpretable/intermediate results, implying that there were none. Bureau *et al.*'s[129] FibroScan study reported the overall number of uninterpretable results, but did not specify how many related to patients with ALD; it is not clear

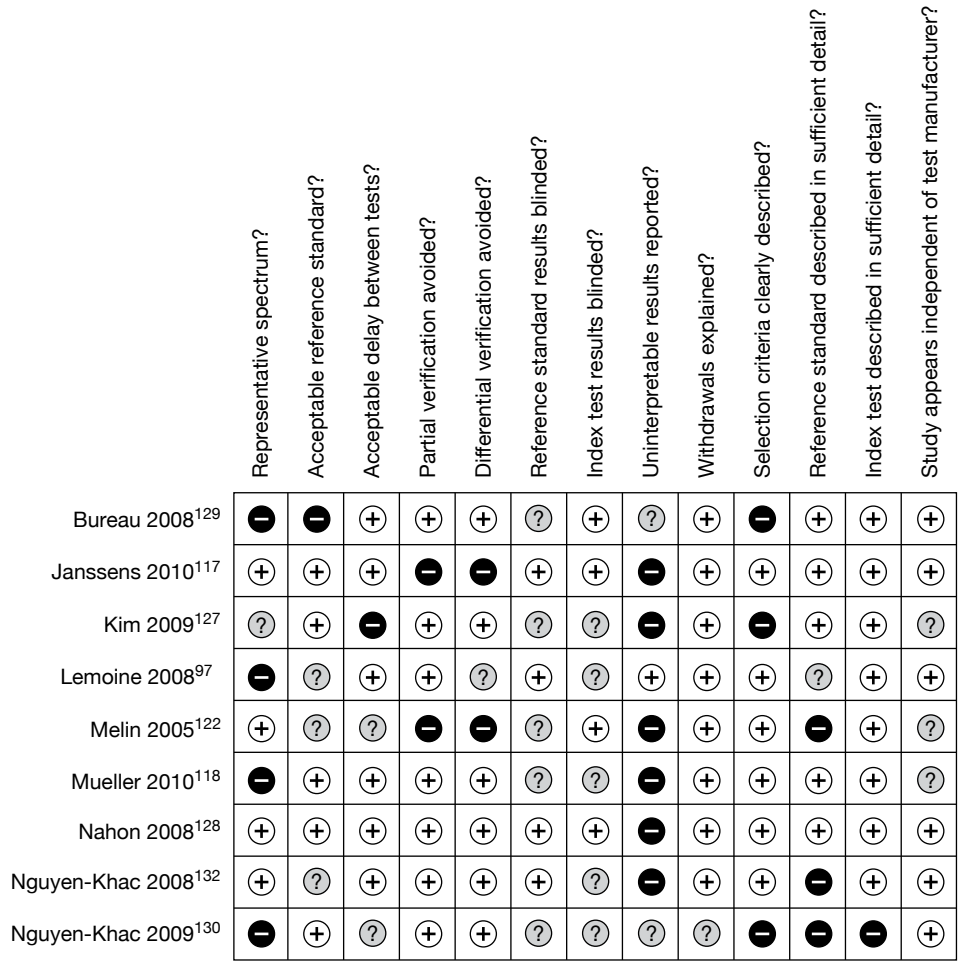



**FIGURE 9** FibroScan: methodological quality summary. Review authors' judgements about each methodological quality item for each included study. –, no; +, yes; ?, unclear.

whether or not they were included in the analyses. Janssens *et al.*,[117] Kim *et al.*,[127] Melin *et al.*,[122] Mueller *et al.*,[118] Nahon *et al.*[128] and Nguyen-Khac *et al.*[132] all reported the number of instances of FibroScan failure (i.e. no or uninterpretable results), but did not include them in their analyses. As none stated how many of these failures occurred in patients who tested positive and how many in patients who tested negative by the reference standard, their impact on test sensitivity and specificity could not be calculated.

The three tests for which evidence has been identified vary in the extent to which that evidence is independent of the test manufacturer. There is no wholly independent evidence relating to the ELF test: one of the investigators, Professor William Rosenberg, is the founder of, and holds stocks in, iQur Ltd, which holds a limited licence to conduct ELF assays on behalf of Siemens Healthcare Diagnostics.[57] Only one of the studies of FibroTest, that by Nguyen-Khac *et al.*,[132] appears to be independent. The remaining studies include in their authorship Thierry Poynard, a major stockholder in, and Mona Munteanu, an employee of, the manufacturers, BioPredictive.[125] However, six of the nine studies that provided data relating to FibroScan[97,117,128–130,132] stated that the authors had no conflicts of interest in relation to the work; of the studies that did not include such a declaration, that by Mueller *et al.*[118] stated that it was funded from independent sources, and only the studies by Kim *et al.*[127] and Melin *et al.*[122] contained no relevant information on this point.

Two further indices of methodological quality proposed by Tsochatzis *et al.*[138] were not included in the QUADAS checklist as they were not applicable to all of the included studies. These were:

- whether studies that used liver biopsy as the reference standard reported that it was performed to an acceptable standard (i.e. the specimen was at least 15 mm long and included at least six portal tracts)
- whether studies of FibroScan reported that it was performed in accordance with the manufacturer's instructions, i.e. using at least 10 valid shots, with a success rate (ratio of valid shots to total number of shots) of at least 60%, and an interquartile range < 30% of the median value.[79]

Only one study that used liver biopsy as a reference standard reported using adequate criteria; this was the study by Janssens *et al.*[117] This was also one of only two studies which clearly stated that FibroScan was performed either in accordance with the manufacturer's instructions or, in the case of the study by Lemoine *et al.*,[97] using more stringent criteria. The remaining studies failed either to meet or, more frequently, to report one or more of the standards (for details, see *Table 14*).

## *Assessment of diagnostic and prognostic accuracy*

### Enhanced Liver Fibrosis Test: diagnostic and prognostic accuracy results

The evidence base for the ELF test is small, resting on a single study of the European Liver Fibrosis Test carried out in a population with chronic liver disease that included < 100 patients diagnosed with ALD. Moreover, the quality of that evidence is not ideal as liver biopsy was not performed to an acceptable standard (see *Study quality*).

This limited evidence suggests that the ELF test can generally distinguish patients with moderate-to-severe fibrosis (Scheuer stages 3–4) from those with milder or no fibrosis (Scheuer stages 0–2) in patients with ALD. Using a low threshold score of 0.087, the test showed 100% sensitivity, but only 16.7% specificity. Ninety-three per cent sensitivity and 100% specificity were achieved using a threshold score of 0.431 (*Table 15*). These threshold scores appear to have been derived from the AUROCs after data collection, rather than prospectively selected and validated. As the the investigators note, because the results rest on data from so few patients, the resulting positive and negative predictive values should be interpreted with caution.[121]

**TABLE 14** FibroScan: reported compliance with manufacturer's instructions

| Study | ≥ 10 valid shots | Success rate > 60% | Interquartile range < 30% of median LSM value |
|---|---|---|---|
| Bureau 2008[129] | Yes | Yes | NR |
| Janssens 2010[117] | Yes | Yes | Yes |
| Kim 2009[127] | Yes | No (≥ 30%) | NR |
| Lemoine 2008[97] | Yes | Yes (≥ 70%) | Yes |
| Melin 2005[122] | No (< 10) | NR | NR |
| Mueller 2010[118] | Yes | Yes | No (< 40%) |
| Nahon 2008,[128] Nahon 2007[131] | Yes | No (≥ 50%) | NR |
| Nguyen-Khac 2008[132] | Yes | Yes | NR |
| Nguyen-Khac 2009[130] | NR | NR | NR |

NR, not reported.

**TABLE 15** Diagnostic and prognostic accuracy of the ELF test in patients with known or suspected ALD

| Condition of interest | Study | Number of assessed patients with known or suspected ALD | Prevalence of condition of interest (%) | Threshold | 2 × 2 data | Sensitivity (%) (95% CI) | Specificity (%) (95% CI) | PPV (%) (95% CI) | NPV (%) (95% CI) | AUROC (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| 'Moderate/severe' fibrosis (Scheuer stages 3–4) | Rosenberg 2004[121] | 64 | *71* | 0.087 | NR | 100 *(97 to 100)* | 16.7 *(7 to 30)* | 75.0 *(67 to 82)* | 100 *(67 to 100)* | 0.94 (0.84 to 1.00) |
| | | | | 0.431 | NR | 93.3 *(87 to 97)* | 100 *(93 to 100)* | 100 *(97 to 100)* | 85.7 *(74 to 94)* | |
| Probable or definite cirrhosis (Scheuer stage 4) | Rosenberg 2004[121] | 64 | NR | NR | NR | NR | NR | NR | NR | 0.83 (0.73 to 0.92) |
| Liver-related clinical outcomes at 6 years | Parkes 2010[57] | 85 | 40 | NR | Not reported | NR | NR | NR | NR | 0.80 (0.70 to 0.98)[a] |
| All-cause mortality at 6 years | Parkes 2010[57] | 85 | 40 | NR | NR | NR | NR | NR | NR | 0.76 (0.65 to 0.82)[a] |

NPV, negative predictive value; NR, not reported; PPV, positive predictive value.

a Professor William Rosenberg, University College London, 2010, personal communication.

Data in roman were taken directly from the text; data in *italics* were calculated by the reviewers.

The ELF test appears to be less successful in distinguishing patients with probable or definite cirrhosis (Scheuer stage 4) from those with milder or no fibrosis (Scheuer stages 0–3) (see *Table 15*). As this result was only presented in the form of an AUROC, the sensitivity and specificity associated with a specific threshold score or scores could not be calculated. Discordant results were not discussed for either fibrosis range.

The evidence relating to the prognostic accuracy of the ELF is derived from data from 85 patients with ALD who were enrolled in the European Liver Fibrosis Test study in English centres and were followed up over a median period of 6.86 years (range 0–9 years). Thus, as for test accuracy, the evidence base is very small. During the follow-up period, 27 patients (32%) died of liver-related causes, a further seven (8%) suffered non-fatal liver-related clinical events, and seven (8%) died of non-liver-related causes.[57] Again, results are only presented in the form of an AUROC. Although this suggests that the ELF is predictive both of liver-related clinical outcomes and of all-cause mortality (for details, see *Table 15*), it should be noted that the sensitivity and specificity associated with a specific threshold score or scores could not be calculated and, more importantly, no information was presented on post-test alcohol consumption, although is likely to have been a substantial confounding factor.

### FibroTest: diagnostic and prognostic accuracy results

The evidence base for FibroTest, although more substantial than that for the ELF test, derives from a total of only 622 patients enrolled in five small to medium-sized studies; although the evidence for test accuracy relative to liver biopsy is derived from a total of only 390 patients enrolled in three of those studies, none of which state that they stipulated a minimum biopsy length of 15 mm.

The available evidence suggests that FibroTest can distinguish between patients with cirrhosis and those with METAVIR stage F0–F3 fibrosis, and, with lesser accuracy, between those with stage F2–F4 and stage F0–F1 fibrosis, and between those with stage F3–F4 and stage F0–F2 fibrosis (*Table 16*). However, not only are these conclusions based on data from only three relatively small studies,[13,123,132] in which some biopsy samples may not have met the recommended minimum standards, as noted above, but the prevalence of the condition of interest was high in each of the three studies, ranging from 63% to 98% for METAVIR stage F2–F4 fibrosis, 51% for METAVIR stage F3–F4 fibrosis, and from 31% to 92% for cirrhosis.

The largest study of test accuracy relative to liver biopsy, that by Naveau *et al.*,[123] also had the most representative population in that it included the lowest proportion of patients with F2–F4 fibrosis. It explored the impact of different threshold scores on sensitivity and specificity. At a threshold score of 0.30, FibroTest had reasonable sensitivity, but rather disappointing specificity in relation to moderate-to-severe (F2–F4) fibrosis; this situation was reversed when the threshold score was raised to 0.70. For cirrhosis, a threshold score of 0.30 produced 100% sensitivity but only 50% specificity, and the balance was improved using a threshold score of 0.70 (for details, see *Table 16*). Thabut *et al.* 2007a[13] found that, at a threshold of 0.48, the specificity of FibroTest in relation to moderate-to-severe (F2–F4) fibrosis was 0% because the prevalence of that condition in the study population, as indicated by liver biopsy, was 98%. Similarly, although FibroTest displayed a sensitivity of 95% for the diagnosis of F2–F4 fibrosis, this result is not robust because 92% of the study population had biopsy results indicative of cirrhosis. The third study of test accuracy relative to liver biopsy, that by Nguyen-Khac *et al.*,[132] did not indicate what diagnostic thresholds were used; it did not report sensitivity and specificity, and the underlying data that would have allowed them to be calculated could not be obtained.

Both Naveau *et al.*[123] and Thabut *et al.* 2007a[13] provided some discussion of discordant cases. Naveau *et al.*[123] reported a discordance of two or more fibrosis stages in 19% of assessed patients

**TABLE 16** Diagnostic and prognostic accuracy of FibroTest in patients with known or suspected ALD

| Condition of interest | Study | Number of assessed patients with known or suspected ALD | Prevalence of condition of interest (%) | Threshold | 2 × 2 data | Sensitivity (%) (95% CI) | Specificity (%) (95% CI) | PPV (%) (95% CI) | NPV (%) (95% CI) | AUROC (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| Mild–severe fibrosis (F1–F4) | Nguyen-Khac 2008[132] | 103 | 92 | NR | NR | NR | NR | NR | NR | 0.77 (0.63 to 0.90) |
| Moderate–severe fibrosis (F2–F4) | Naveau 2005[123] | 221 | 63 | 0.30 | TP: 125; FP: 28; FN: 23; TN: 54[a] | 84 *(78 to 90)* | 66 *(55 to 76)* | 82 *(75 to 87)* | 70 *(60 to 80)* | 0.83 (0.81 to 0.87) |
| | | | | 0.70 | TP: 81; FP: 6; FN: 67; TN: 76[a] | 55 *(47 to 63)* | 93 *(86 to 97)* | 93 *(87 to 97)* | 53 *(45 to 61)* | |
| | Nguyen-Khac 2008[132] | 103 | 75 | NR | NR | NR | NR | NR | NR | 0.79 (0.69 to 0.90) |
| | Thabut 2007a[13] | 66 | 98 | 0.48[b] | TP: 62; FP: 1; FN: 3; TN: 0[b] | *95*[c] *(88 to 99)* | *0*[c] *(0 to 85)* | *98*[c] *(93 to 100)* | *0*[c] *(0 to 54)* | 0.83- ±–0.03[b] |
| Severe fibrosis (F3–F4) | Nguyen-Khac 2008[132] | 103 | 51 | NR | NR | NR | NR | NR | NR | 0.80 (0.70 to 0.91) |
| Cirrhosis (F4) | Naveau 2005[123] | 221 | 31 | 0.30 | TP: 68; FP: 76; FN: 0; TN: 77 | 100 *(96 to 100)* | 50 *(42 to 58)* | 47 *(39 to 55)* | 100 *(97 to 100)* | 0.95 (0.94 to 0.96) |
| | | | | 0.70 | TP: 62; FP: 20; FN: 6; TN: 133 | 91 *(83 to 97)* | 87 *(81 to 92)* | 76 *(66 to 84)* | 96 *(90 to 98)* | |
| | Nguyen-Khac 2008[132] | 103 | 32 | NR | NR | NR | NR | NR | NR | 0.84 (0.72 to 0.97) |
| | Thabut 2007a[13] | 66 | 92 | 0.74[b] | TP: 47; FP: 1; FN: 13; TN: 5[b] | *78*[c] *(66 to 87)* | *83*[c] *(44 to 98)* | *98*[c] *(91 to 100)* | *28*[c] *(10 to 50)* | 0.90 ± 0.04[b] |
| HVPG ≥ 5 mmHg | Thabut 2007a[13] | 66 | *94*[c] | 0.48[b] | TP: 48; FP: 1; FN: 14; TN: 3[b] | *77*[c] *(65 to 87)* | *75*[c] *(28 to 97)* | *98*[c] *(91 to 100)* | *18*[c] *(5 to 40)* | 0.92 ± 0.03[b] |
| HVPG ≥ 12 mmHg | Thabut 2007a[13] | 66 | *86*[c] | 0.58[b] | TP: 53; FP: 2; FN: 4; TN: 7[b] | *93*[c] *(84 to 98)* | *87*[c] *(59 to 99)* | *98*[c] *(92 to 100)* | *67*[c] *(39 to 88)* | 0.84 ± 0.03[b] |

*continued*

**TABLE 16** Diagnostic and prognostic accuracy of FibroTest in patients with known or suspected ALD *(continued)*

| Condition of interest | Study | Number of assessed patients with known or suspected ALD | Prevalence of condition of interest (%) | Threshold | 2 × 2 data | Sensitivity (%) (95% CI) | Specificity (%) (95% CI) | PPV (%) (95% CI) | NPV (%) (95% CI) | AUROC (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| Oesophageal varices (grade 2) | Thabut 2003[125] | 58[b] | *76*[c] | 0.85[b] | TP: 39; FP: 7; FN: 5; TN: 7[b] | *89*[c] *(78 to 96)* | *50*[c] *(25 to 77)* | *85*[c] *(73 to 94)* | *58*[c] *(31 to 83)* | 0.73 ± 0.08[b] |
| All-cause mortality at 2 months | Thabut 2007b[126] | 189 | 39[b] | NR | NR | NR | NR | NR | NR | 0.64 ± 0.05[b] |
| All-cause mortality at 6 months | Thabut 2007b[126] | 189 | 54 | NR | NR | NR | NR | NR | NR | 0.58 ± 0.05[b] |
| Liver-related mortality at 5 years | Naveau 2009[124] | 218 | 19 | ≥ 0.31 | NR | NR | NR | NR | NR | 0.79 (0.68 to 0.86) |
| | | | | 0.33–0.58 | NR | NR | NR | NR | NR | |
| | | | | ≥ 0.59 | NR | NR | NR | NR | NR | |
| All-cause mortality at 5 years | Naveau 2009[124] | 218 | 39 | ≥ 0.31 | NR | NR | NR | NR | NR | 0.69 (0.61 to 0.76) |
| | | | | 0.33–0.58 | NR | NR | NR | NR | NR | |
| | | | | ≥ 0.59 | NR | NR | NR | NR | NR | |

NPV, negative predictive value; NR, not reported, PPV, positive perdictive value.
a It is not clear why these figures total 230, when only 221 patients were included in the study.
b Dr Thierry Poynard, Hôpital Pitié-Salpêtriére, Paris, 2010, personal communication.
c Calculated from data provided by Dr Thierry Poynard, Hôpital Pitié-Salpêtriére, Paris, 2010, personal communication.
Data in roman were taken directly from the text; data in *italics* were calculated by the reviewers.

(42/221). On the basis of independent clinical, ultrasonographical, and endoscopic signs of cirrhosis, they attributed the error to the biopsy in 26 cases (14 FNs and 12 FPs), and to FibroTest in 13 cases (six FNs and seven FPs); three cases were unattributable.[123] Eighteen of the 42 discordant cases involved diagnoses of cirrhosis: three FNs and three possible FPs of FibroTest, three FPs and eight possible FNs of biopsy (the eight FNs of biopsy were all in poor-quality samples), and one unattributable case diagnosed as cirrhosis by FibroTest but not by biopsy.[123]

Thabut *et al.* 2007a[13] attributed discordant cases to failure of biopsy or FibroTest on the basis of clinical events (haemorrhage, ascites) and risk factors for FibroTest failure. Four of the 61 patients with cirrhosis (7%) had FN results on FibroTest; all had large ascites and low alpha-2-macroglobulin. No other discordant results were reported.[13]

Two small studies by Thabut *et al.* (2007a and 2003),[13,125] which included only 66 and 58 patients with ALD, respectively, suggest that FibroTest can also distinguish between patients with and without PHt and, with less accuracy, between those with and without oesophageal varices. However, these studies were also carried out in populations with a high prevalence of those conditions, and indeed the investigators noted that, as 86% of the population of Thabut *et al.*'s 2007a study[13] had HVPG results indicating clinically significant PHt (HVPG > 12 mmHg), the study findings should not be used as a basis for recommending the use of FibroTest alone to predict severe PHt in cirrhotic patients.

The study by Naveau *et al.*[124] and, to a lesser extent, that by Thabut *et al.* 2007b[126] suggest that FibroTest may also be able, with relatively low accuracy, to predict liver-related mortality and all-cause mortality (see *Table 16*). In Naveau *et al.*'s[124] cohort study, 85 patients (39%) died during the follow-up period: 42 (19%) of liver-related causes (haemorrhage, HCC, and decompensation) and 43 (20%) of non-liver-related causes. FibroTest was predictive of survival or non-liver-related mortality and, to a lesser degree, of overall mortality (for details, see *Table 16*). Details of 5- and 10-year survival according to baseline FibroTest values are presented in *Tables 17* and *18*. The baseline FibroTest and biopsy results were concordant for 38 (90%) of the 42 liver-related deaths (29 with cirrhosis, nine without cirrhosis) and discordant for only four (10%; two FPs of FibroTest, and one FP and one FN of biopsy).[124]

Naveau *et al.*'s[124] cohort study also provided information on subsequent alcohol consumption in patients enrolled in their 2005 study of test accuracy.[124] Only 21% (46/218) were known to be abstinent during the follow-up period; 50% (108/218) were not abstinent, and the status of the remaining 29% (64/218) was not known.[124] Unfortunately, the authors did not link these data with test results, and thus it was not possible to determine whether or not the test results had an impact on subsequent alcohol consumption, or whether or not alcohol consumption affected survival.

### FibroScan: diagnostic accuracy results

The evidence base for FibroScan is slightly larger than that for FibroTest, deriving from a total of approximately 868 patients enrolled in nine small to medium-sized studies (the total number of participants is approximate because, as indicated in *Table 13*, it is not always clear how many of the eligible patients were in fact assessed). The evidence for test accuracy relative to liver biopsy is derived from a total of only 480 patients enrolled in the studies by Janssens *et al.*,[117] Kim *et al.*,[127] Melin *et al.*,[122] Mueller *et al.*,[118] Nahon *et al.*,[128] and Nguyen-Khac *et al.*[132] In only one of these, that by Janssens *et al.*,[117] are the liver biopsy samples known to have met the recommended minimum standards.

The evidence that FibroScan can distinguish patients with METAVIR stage F1–F4 fibrosis from those without fibrosis (F0), and those with F2–F4 fibrosis from those with stage F0–F1, is not

**TABLE 17** Five-year survival in patients with ALD, by baseline FibroTest value (after Naveau S, Gaudé G, Asnacios A, Agostini H, Abella A, Barri-Ova N, *et al*. Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology* 2009;**49**:97–105.) Reproduced with permission from John Wiley & Sons.

| Baseline FT value | *n* | Liver-related death | Survival or non-liver-related death (%) (95% CI) | Death, any cause | Overall survival (%) (95% CI) |
|---|---|---|---|---|---|
| 0.00–0.31 | 81 | 1 | 98.7 (96.0 to 100) | 9 | 88.9 (82.0 to 95.7) |
| 0.32–0.58 | 43 | 3 | 92.1 (83.5 to 100) | 7 | 83.4 (72.1 to 94.6) |
| 0.59–1.00 | 94 | 28 | 68.3 (58.5 to 78.0) | 39 | 58.4 (48.4 to 68.4) |
| All | 218 | 32 | 84.5 (79.5 to 89.4) | 55 | 74.7 (68.9 to 80.5) |

FT, FibroTest.

**TABLE 18** Ten-year survival in patients with ALD, by baseline FibroTest value (after Naveau S, Gaudé G, Asnacios A, Agostini H, Abella A, Barri-Ova N, *et al*. Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology* 2009;**49**:97–105) Reproduced with permission from John Wiley & Sons.

| Baseline FT value | *n* | Liver-related death | Survival or non-liver-related death: (%) (95% CI) | Death, any cause | Overall survival (%) (95% CI) |
|---|---|---|---|---|---|
| 0.00–0.31 | 81 | 5 | 92.0 (84.9 to 99.0) | 21 | 71.4 (60.7 to 82.2) |
| 0.32–0.58 | 43 | 4 | 87.5 (75.5 to 99.5) | 12 | 69.8 (55.2 to 84.4) |
| 0.59–1.00 | 94 | 32 | 62.6 (52.2 to 73.1) | 50 | 42.4 (31.1 to 53.6) |
| All | 218 | 41 | 78.5 (72.4 to 84.6) | 83 | 58.8 (51.7 to 65.9) |

FT, FibroTest.

robust, being based on only one fairly small study by Nguyen-Khac *et al.*[132] in a population with a high prevalence of stage F2–F4 fibrosis. However, there is more substantial evidence that FibroScan can distinguish patients with stage F3–F4 from those with stage F0–F2 fibrosis, and those with cirrhosis from those with stage F0–F3 fibrosis (*Table 19*).

The data presented in *Table 19* illustrate the impact of different threshold scores on the sensitivity and the specificity of FibroScan. Castéra *et al.*[81] originally suggested that, in patients with chronic hepatitis C, the optimal FibroScan threshold values for the identification of significant (F2–F4) fibrosis, advanced (F3–F4) fibrosis, and cirrhosis (F4) were 7.1, 9.5, and 12.5 kPa, respectively (*Table 20*). These values were used by some of the studies included in this review and were found to be less appropriate for use in patients with ALD. Melin *et al.*[122] achieved 100% sensitivity using a threshold score of 13 kPa to identify cirrhosis in patients with ALD, but could not estimate the specificity associated with this threshold because patients with a score < 13 kPa did not undergo liver biopsy. Other investigators specifically sought to identify the optimal threshold scores for use in patients with ALD (see *Table 19*). Janssens *et al.*[117] noted that, in such patients, the threshold score of 9.5 kPa proposed for hepatitis C had 100% sensitivity for identifying severe (F3–F4) fibrosis, but a PPV of only 65%; it overestimated the degree of fibrosis in 17 of the 49 patients (35%) who underwent liver biopsy, and in all but one of them did so by two or more stages. Janssens *et al.*,[117] therefore, suggested that more appropriate thresholds for the identification of severe (F3–F4) fibrosis in ALD would lie between 15.8 and 17.3 kPa. They did not report the sensitivity and the specificity of a threshold score of 12.5 kPa for identifying cirrhosis (F4) in patients with ALD, but suggested that a more appropriate threshold would lie between 19.6 and 23.5 kPa, the exact choice depending on the preferred balance between sensitivity and specificity.[117] This study was not ideally designed to establish specificity because biopsy was offered only to patients with a FibroScan score of ≥ 9.5 kPa.

**TABLE 19** Diagnostic accuracy of FibroScan in patients with known or suspected ALD

| Condition of interest | Study | Number of assessed patients with known or suspected ALD | Prevalence of condition of interest, % | Threshold (kPa) | 2 × 2 data | Sensitivity, % (95% CI) | Specificity, % (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) | AUROC (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| Mild–severe fibrosis (F1–F4) | Nguyen-Khac 2008[132] | 103 | 92 | 5.9 | NR | 83 *(75 to 90)* | 86 *(55 to 99)* | *99*[a] *(94 to 100)* | *35*[a] *(14 to 50)* | 0.84 (0.73 to 0.95) |
| Moderate–severe fibrosis (F2–F4) | Nguyen-Khac 2008[132] | 103 | 75 | 7.8 | NR | 80 *(70 to 88)* | 90.5 *(78 to 98)* | 97 *(90 to 99)* | 60 *(45 to 74)* | 0.91 (0.85 to 0.97) |
| Severe fibrosis (F3–F4) | Janssens 2009,[119] 2010[117] | 49 | 65 | 9.5 | TP: 32; FP 17; FN: 0; TN: 0 | *100 (93 to 100)* | *0 (0 to 14)* | 65 *(52 to 78)* | *Not calculable (0 to 100)* | 0.77 |
| | | | | 13.9 | NR | 81 *(66 to 93)* | 59 *(35 to 80)* | *79 (64 to 91)* | *63 (38 to 84)* | |
| | | | | 15.8 | NR | 75 *(59 to 88)* | 70 *(48 to 89)* | *83 (67 to 94)* | *60 (38 to 80)* | |
| | | | | 16.5 | NR | *72 (55 to 86)* | 70 *(48 to 89)* | *82 (66 to 94)* | *57 (36 to 77)* | |
| | | | | 17.0 | NR | *72 (55 to 86)* | *77 (53 to 91)* | *85 (69 to 95)* | *59 (38 to 78)* | |
| | | | | 17.3 | NR | 69 *(52 to 83)* | *77 (53 to 91)* | *85 (67 to 95)* | *57 (36 to 76)* | |
| | Kim 2009[127] | 45 | 80 | NR | TP: 19; FP: 1; FN: 10; TN: 15 | *66 (48 to 81)* | *94 (74 to 99)* | *95 (79 to 99)* | *60 (41 to 78)* | 0.98 (0.94 to 1.0) |
| | Mueller 2010[118] | 101 | 45 | 8.0 | NR | 91 *(80 to 97)* | 75 *(63 to 85)* | *75 (62 to 85)* | *91 (81 to 97)* | 0.91 ± 0.03 |
| | Nahon 2008[128] | 147 | 75 | 11.6 | NR | 87 *(80 to 93)* | 89 *(76 to 96)* | 96 *(91 to 99)* | 70 *(57 to 82)* | 0.94 (0.90 to 0.97) |
| | Nguyen-Khac 2008[132] | 103 | 51 | 11.0 | NR | 87 *(76 to 94)* | 80.5 *(69 to 90)* | 82 *(71 to 91)* | 84 *(74 to 94)* | 0.90 (0.82 to 0.97) |

*continued*

**TABLE 19** Diagnostic accuracy of FibroScan in patients with known or suspected ALD (*continued*)

| Condition of interest | Study | Number of assessed patients with known or suspected ALD | Prevalence of condition of interest, % | Threshold (kPa) | 2 × 2 data | Sensitivity, % (95% CI) | Specificity, % (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) | AUROC (95% CI) |
|---|---|---|---|---|---|---|---|---|---|---|
| Cirrhosis (F4) | Janssens 2010[117] | 49 | 41 | 12.5 | NR | NR | NR | NR | NR | 0.86 |
| | | | | 19.6 | NR | 80 *(59 to 93)* | 76 *(59 to 89)* | *70 (50 to 86)* | *85 (67 to 95)* | |
| | | | | 21.1 | NR | 75 *(54 to 91)* | 80 *(63 to 92)* | *71 (51 to 88)* | *82 (66 to 94)* | |
| | | | | 23.5 | NR | 65 *(43 to 84)* | 83 *(67 to 94)* | *72 (50 to 90)* | *77 (61 to 90)* | |
| | Kim 2009[127] | 45 | 64 | 28.5 | NR | 90 *(75 to 97)* | 87 *(66 to 97)* | *93 (79 to 98)* | *82 (60 to 95)* | 0.97 (0.93 to 1.0) |
| | Melin 2005[122] | 35 | 97 | 13.0 | TP: 34; FP: 1; FN: 0; TN: 0 | *100 (93 to 100)* | *0 (0 to 85)* | *97 (87 to 100)* | Not calculable *(0 to 100)* | NR |
| | Mueller 2010[118] | 101 | 26 | 11.5 | NR | 100 *(91 to 100)* | 77 *(67 to 86)* | *60 (46 to 74)* | *100 (96 to 100)* | 0.92 (0.87 to 0.97) |
| | | | | 12.5 | TP:25; FP: 15; FN: 1; TN: 60[b] | 96 *(83 to 100)* | 80 *(70 to 88)* | 63[b] *(47 to 77)* | 98[b] *(93 to 100)* | |
| | Nahon 2008[128] | 147 | 54 | 22.7 | NR | 84 *(75 to 91)* | 83 *(74 to 91)* | *86 (77 to 93)* | *81 (71 to 89)* | 0.87 (0.81 to 0.93) |
| | Nguyen-Khac 2008[132] | 103 | 32 | 12.8 | NR | 100 *(93 to 100)* | 75 *(64 to 84)* | *65 (51 to 77)* | *100 (95 to 100)* | 0.92 (0.87 to 0.98) |
| | | | | 19.5 | NR | 86 *(71 to 95)* | 84 *(75 to 92)* | *72 (57 to 85)* | *92 (85 to 97)* | |
| HVPG ≥ 10 mmHg | Bureau 2008[129] | ≤ 51 | 51 | 21.0 | NR | NR | NR | NR | NR | 0.91 (0.76 to 1.01) |
| | Janssens 2010[117] | 49 | Not clear | 17.0 | NR | 80 | 55 | NR | NR | 0.81 |
| | | | | 24.7 | NR | 67 | 80 | NR | NR | |
| | Lemoine 2008[97] | 48 | 83 | 34.9 | NR | 90 *(78 to 97)* | 88 *(55 to 99)* | 97 *(88 to 100)* | 64 *(35 to 86)* | 0.94 ± 0.03 |
| HVPG ≥ 12 mmHg | Janssens 2010[117] | 49 | Not clear | 25.7 | NR | 75 | 80 | NR | NR | 0.80 |
| Large oesophageal varices | Nguyen-Khac 2009[130] | 103 | *25* | 47.2 | NR | 85 *(67 to 95)* | 64 *(53 to 74)* | *41 (13 to 58)* | *94 (83 to 97)* | 0.77 (0.68 to 0.85) |

NPV, negative predictive value; NR, not reported, PPV, positive predictive value.

a PPV and NPV were reported by Nguyen-Khac as 98% and 35%, respectively, but have been recalculated here as these figures were inconsistent with the reported sensitivity and specificity, which have been assumed to be correct.

b Professor Sebastian Mueller, University of Heidelberg, Germany, 2010, personal communication.

Data in roman were taken directly from the text; data in *italics* were calculated by the reviewers.

Three studies discussed discordant results.[97,117,128] Nahon *et al.*[128] found that, relative to biopsy, FibroScan underestimated and overestimated the degree of fibrosis in approximately equal proportions of patients. Fourteen per cent (11/79) of those with histologically proven cirrhosis had FibroScan scores < 22.6 kPa, whereas 16% (11/68) of those with FibroScan scores > 22.6 kPa had biopsy results that did not indicate cirrhosis, although the majority (10/11) displayed extensive fibrosis. Janssens *et al.*[117] found that FibroScan overestimated the degree of fibrosis in 7 of the 11 patients with severe steatosis (by two stages in five patients and by one stage in two patients). Of the six patients in the study with alcoholic hepatitis, FibroScan classified three as having cirrhosis, although their biopsy results indicated F3 fibrosis. In two of the remaining three, both FibroScan and biopsy results indicated cirrhosis, thus removing the potential for overestimation by FibroScan. Finally, Lemoine *et al.*[97] noted one discordant case in a 70-year-old patient who had been totally abstinent from alcohol for 12 months and had no histological indication of alcoholic hepatitis. The patient's FibroScan score was 38.10 ± 10 kPa, suggestive of PHt, although his HVPG measurement was only 8 mmHg.

Like Janssens *et al.*,[117] Mueller *et al.*[118] found that in patients with inflammatory hepatitis, liver stiffness was increased independently of the degree of fibrosis. They, therefore, found that the diagnostic accuracy of FibroScan improved when patients with laboratory signs of alcoholic steatohepatitis [i.e. serum glutamic oxaloacetic transaminase (SGOT) levels above 100 units/litre (U/L)] were excluded. When patients with only mildly elevated SGOT (> 50 U/L) were excluded, diagnostic accuracy improved in relation to F3–F4 fibrosis, but not in relation to F4 (cirrhosis) (*Table 21*).

Three small studies[97,117,129] suggest that FibroScan can generally distinguish between patients with and without PHt, whereas one smallish study[130] suggests that it can distinguish with less success between patients with and without large oesophageal varices (see *Table 19*).

**TABLE 20** Optimal liver stiffness cut-off values for the diagnosis of fibrosis in patients with chronic hepatitis C (from Castéra *et al.*[81])

| Degree of fibrosis | Optimal cut-off (kPa) | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|---|---|---|---|---|---|
| F2–F4 | 7.1 | 67 | 89 | 95 | 48 |
| F3–F4 | 9.5 | 73 | 91 | 87 | 81 |
| F4 | 12.5 | 87 | 91 | 77 | 95 |

NPV, negative predictive value; PPV, positive predictive value.

**TABLE 21** Diagnostic accuracy of FibroScan in patients with suspected ALD, with and without those with elevated SGOT (from Mueller *et al.*[118])

| Disease severity and SGOT status | Included patients (*n*) | AUROC | Standard error | Threshold value (kPa) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|---|
| F4 | 101 | 0.921 | 0.03 | 11.5 | 100 | 77 |
| | | | | 12.5 | 96 | 80 |
| F4 without SGOT > 100 U/L | 86 | 0.944 | 0.02 | 11.5 | 100 | 84 |
| | | | | 12.5 | 95 | 90 |
| F4 without SGOT > 50 U/L | 66 | 0.945 | 0.03 | 10.4 | 100 | 87 |
| | | | | 12.5 | 92 | 91 |
| F3–F4 | 101 | 0.914 | 0.03 | 8.0 | 91 | 75 |
| F3–F4 without SGOT > 100 U/L | 80 | 0.922 | 0.03 | 8.0 | 87 | 87 |
| F3–F4 without SGOT > 50 U/L | 67 | 0.946 | 0.03 | 8.0 | 100 | 84 |

There are no long-term data relating FibroScan results to survival or other clinical outcomes.

## *Adverse events and failure rates*

### The Enhanced Liver Fibrosis Test and FibroTest: failure rates and adverse events

Both the ELF test and FibroTest use blood samples obtained by standard venepuncture. None of the included studies reported adverse events relating to this process. However, a systematic review of studies of adverse events in adults undergoing simple venepuncture for diagnostic or screening purposes indicates that between 14% and 45% of patients undergoing such venepuncture suffer pain and bruising, whereas between 0.9% and 3.4% suffer vasovagal reactions. Potentially disabling nerve injuries occur but, fortunately, appear to be very rare (for full details, see *Appendix 8*).

There are no data relating to test failure rates for the ELF test or FibroTest specifically in patients with ALD. The only relevant data come from Naveau *et al.*'s study[123] of FibroTest in which, for unspecified reasons, serum samples were unavailable for 17% (50/292) of the enrolled patients. In Rosenberg *et al.*'s study[121] of the ELF test, 4.4% (45/1021) of patients overall had incomplete clinical details or biochemical samples, compared with 5.6% (55/976) whose biopsy samples were considered inadequate; figures relating specifically to patients with ALD were not presented.

### FibroScan: failure rates and adverse events

Only two of the included studies commented on the acceptability of FibroScan to patients.[117,132] Janssens *et al.*[117] found that only 2% (5/255) of patients entering hospital for alcohol detoxification and rehabilitation refused FibroScan, whereas 29% (21/72) of those with FibroScan results indicative of severe fibrosis or cirrhosis-refused liver biopsy. However, Nguyen-Khac *et al.*[132] found that 34% (55/160) of patients refused to participate in their study, which involved both transient elastography and venepuncture; it is not clear whether this reluctance to participate related specifically to one or other of those interventions, or to the perceived inconvenience of undergoing both.

Some studies reported the proportion of patients with ALD in whom FibroScan was unsuccessful: this ranged between 4.4% and 8.6% (*Table 22*). Other studies reported the number of potential participants who were excluded because of either failure to obtain a valid result or the presence of obesity or other factors likely to affect FibroScan performance. In the study by Kim *et al.*,[127] 11.8% of patients were excluded because of the factors likely to affect test performance.

**TABLE 22** Proportion of patients with ALD in whom FibroScan was either thought inappropriate or was unsuccessful

| Study | Exclusion criteria that might affect FibroScan performance | FibroScan unsuccessful: patients with ALD (%) | Reasons for failure |
|---|---|---|---|
| Janssens 2010[117] | None reported | 11/250 (4.4) | Obesity or ascites |
| Kim 2009[127] | BMI ≥ 30; probability of successful testing < 30%: 6/51 (11.8%) excluded for these reasons | None reported | None reported |
| Lemoine 2008[97] | None reported | None reported | None reported |
| Melin 2005[122] | None reported | 18/245 (7.3) | Obesity |
| Mueller 2010[118] | None reported | 5/106 (4.7) | Measurements invalid or interquartile range > 40% |
| Nahon 2008[128] | None reported | 15/174 (8.6) | Results inadequate; no reason given |
| Nguyen-Khac 2008[132] | Failure to obtain valid measurements: 2/105 (1.9%) excluded, no more specific reason given | None reported | None reported |
| Nguyen-Khac 2009[130] | None reported | None reported | None reported |

Although obesity is the most frequently reported reason for FibroScan failure, Bureau *et al.*[129] found that one-third of test failures (2/6) in patients with liver disease of varied aetiology could not be attributed to obesity; their cause remained unclear.

Two studies that sought specifically to identify the features associated with successful FibroScan use did not meet the review's inclusion criteria, but nonetheless provide useful information in this context.[79,139] The largest prospective study of this nature, by Castéra *et al.*,[79] assessed 13,369 examinations performed by seven operators over a 5-year period in 7261 adult patients with chronic liver disease of varied aetiology. It recorded the prevalence of failure of LSM (defined, in accordance with the manufacturer's recommendation, as failure to obtain any value after at least 10 shots) and unreliable results (defined, again in accordance with the manufacturer's recommendations, as < 10 successful shots, a success rate < 60%, or an interquartile range > 30% of the median value). In 18.4% of examinations (2466/13,369), valid results could not be obtained. LSM failed in 3.1% (420/13,369), whereas in 15.8% (2046/12,949) of the remaining examinations the results were deemed unreliable. Although the number of patients in whom FibroScan could be used successfully could be increased with repeated examinations, there remained some for whom it was impossible to obtain either any result or a reliable result after five attempts (*Table 23*).[79]

Castéra *et al.*[79] found that the factor most strongly associated with both test failure and unreliable results at the first FibroScan examination was a BMI > 30 kg/m$^2$ (*Table 24*); the rates of both failure and unreliable results increased in parallel with the BMI (*Table 25*). The rates of both failure and unreliable results were also substantially raised when the operator had performed < 500 examinations (*Table 26*). Such a high threshold was chosen to define operator experience because all the operators who participated in the study had already performed at least 100 FibroScan examinations. The failure rate ranged from 0.2% in lean young non-diabetic patients to 20.9% in elderly obese diabetic patients, while the rate of unreliable results ranged from 7.2% in lean young men without diabetes or hypertension to 60.4% in elderly obese women with diabetes and hypertension.

In a smaller study, Kettaneh *et al.*[139] failed to achieve the manufacturer's recommendation of at least 10 successful LSMs in 8.4% (79/935) of patients with chronic hepatitis C. They found that success was directly related to increased operator experience, and inversely related to both patient age and patient BMI. However, with hindsight, they suggested that the limiting factor was not so much BMI per se as the presence of a fatty thoracic belt that made it technically impossible to obtain accurate results.

A pilot study conducted specifically in patients with a BMI of ≥ 30 kg/m$^2$ found that the use of the XL specialised probe, which can measure to a depth of 35–75 mm below the skin surface, reduced rates of failure and unreliable results. However, even with this probe, no value could be obtained in 12% of patients and the recommended standard of at least 10 valid measurements

**TABLE 23** Effect of repeated examinations on rate of failure of LSM or unreliable results (data from Castéra *et al.*[79])

| Number of examinations undergone by patient | Percentage of patients with | |
|---|---|---|
| | LSM failure | Unreliable results |
| 1 | 4.0 | 17.0 |
| 2 | 2.4 | 15.2 |
| 3 | 2.2 | 14.5 |
| 4 | 1.2 | 14.3 |
| 5 | 1.2 | 9.6 |

**TABLE 24** Factors independently associated with failure of LSM or unreliable results at first FibroScan examination (data from Castéra *et al*.[79])

| Factor | Odds ratio | 95% CI | *p*-value |
|---|---|---|---|
| ***Failure of LSM*** | | | |
| BMI > 30 kg/m$^2$ | 7.5 | 5.6 to 10.2 | 0.0001 |
| Operator experience < 500 examinations | 2.5 | 1.6 to 4.0 | 0.0001 |
| Age > 52 years | 2.3 | 1.6 to 3.2 | 0.0001 |
| Type 2 diabetes (fasting serum glucose > 5.6 mmol/L or ongoing anti-diabetic treatment) | 1.6 | 1.1 to 2.2 | 0.009 |
| ***Unreliable results*** | | | |
| BMI > 30 kg/m$^2$ | 3.3 | 2.8 to 4.0 | 0.0001 |
| Operator experience < 500 examinations | 3.1 | 2.4 to 3.9 | 0.0001 |
| Age > 52 years | 1.8 | 1.6 to 2.1 | 0.0001 |
| Female gender | 1.4 | 1.2 to 1.6 | 0.0001 |
| Hypertension (defined as ongoing hypertensive pharmacological treatment) | 1.3 | 1.1 to 1.5 | 0.003 |
| Type 2 diabetes (fasting serum glucose > 5.6 mmol/L or ongoing anti-diabetic treatment) | 1.2 | 1.0 to 1.5 | 0.05 |

**TABLE 25** Rates of LSM failure and unreliable results at first FibroScan examination, by BMI (data from Castéra *et al*.[79])

| BMI (kg/m$^2$) | LSM failure (%) (*n*) | Unreliable results (%) (*n*) |
|---|---|---|
| < 25 | 1.0 (4172) | 12.0 (4130) |
| ≥ 25 | 8.1 (3089) | 24.3 (2838) |
| ≥ 28 | 12.4 (1568) | 31.2 (1373) |
| ≥ 30 | 16.9 (967) | 35.4 (804) |
| ≥ 35 | 24.9 (225) | 39.1 (169) |
| ≥ 40 | 41.7 (40) | 53.6 (28) |

**TABLE 26** Rates of LSM failure and unreliable results, by operator experience (data from Castéra *et al*.[79])

| Operator experience | LSM failure | Unreliable results |
|---|---|---|
| < 500 examinations (%) | 8.3 | 30.5 |
| > 500 examinations (%) | 3.5 | 15.6 |
| *p*-value | < 0.0001 | < 0.0001 |

could be achieved in only 76%.[140] These findings are particularly important in the light of the high proportion of the UK population with a BMI of ≥ 30 kg/m$^2$ or a raised waist circumference (see *Chapter 1, Incidence and prevalence*).

## *Discussion of clinical effectiveness*

### Diagnostic and prognostic accuracy

#### *Summary of diagnostic and prognostic accuracy of the Enhanced Liver Fibrosis Test*

No studies were identified that specifically assessed the ELF test. One study[121] ($n = 64$ patients) evaluated the diagnostic accuracy of the European Liver Fibrosis Test (essentially the ELF test with the addition of age to the algorithm) relative to liver biopsy, in identifying

moderate-to-severe fibrosis and cirrhosis in patients with known or suspected ALD. The study found that, at a threshold of 0.431, the ELF test identified moderate-to-severe fibrosis with a sensitivity of 93% (95% CI 87% to 97%) and a specificity of 100% (95% CI 93% to 100%). The sensitivity and specificity for cirrhosis were not reported, but were presumably lower: the point estimate of the AUROC for cirrhosis was lower than that for moderate-to-severe fibrosis (0.83 vs 0.94), although the CIs overlapped. As the evidence base is very small, and acceptable minimum standards were not used for the biopsy sample, these findings are not robust.

A follow-up study[57] ($n = 85$ patients) suggested that the test had a predictive value in relation to both liver-related clinical outcomes and all-cause mortality, but did not report sensitivity and specificity. Again, the findings are not robust because the evidence base is so small.

### Summary of diagnostic and prognostic accuracy of FibroTest

Three studies[13,123,132] assessed the diagnostic and prognostic accuracy of FibroTest in identifying moderate-to-severe fibrosis and cirrhosis in patients with known or suspected ALD ($n = 390$ patients). The largest of these studies, that by Naveau *et al.*[123] ($n = 221$ patients), also had the most representative population. The study found that, using a threshold score of 0.30, FibroTest could identify moderate-to-severe (F2–F4) fibrosis with a sensitivity of 84% (95% CI 78% to 90%) and a specificity of 66% (95% CI 55% to 76%), whereas, using a threshold score of 0.70, it could identify cirrhosis with a sensitivity of 91% (95% CI 83% to 97%) and a specificity of 87% (95% CI 81% to 92%). Evidence for FibroTest's ability to distinguish between patients with and without fibrosis (F1–4 vs F0) is not robust, being based on only one fairly small study by Nguyen-Khac *et al.*[132] ($n = 103$ patients) in a population in whom the prevalence of fibrosis or cirrhosis was 92%.

A small study by Thabut *et al.* 2007a[13] ($n = 66$ patients) found that FibroTest could identify clinically significant PHt (HVPG $\geq 12$ mmHg) with a sensitivity of 93% (95% CI 84% to 98%) and specificity of 87% (95% CI 55% to 99%). However, because of the high prevalence of the condition in the study population, the investigators felt that this finding should not be used to support the use of FibroTest alone to predict severe PHt in cirrhotic patients.

A second small study by Thabut *et al.* 2003[125] ($n = 58$ patients) found that, using a threshold of 0.85, FibroTest could predict the presence of grade 2 oesophageal varices with a sensitivity of 89% (95% CI 73% to 96%) and a specificity of 50% (95% CI 25% to 77%).

Finally, a study by Thabut *et al.* 2007b[126] ($n = 189$ patients) suggested that FibroTest has a modest predictive value in relation to all-cause mortality at 2 and 6 months (AUROCs $0.64 \pm 0.05$ and $0.58 \pm 0.05$, respectively); sensitivity and specificity were not reported. A study by Naveau *et al.*[124] ($n = 218$ patients) suggested that FibroTest had a somewhat better predictive value in relation to both liver-related and all-cause mortality at 5 years [AUROCs 0.79 (95% CI 0.68 to 0.86) and 0.69 (95% CI 0.61 to 0.76), respectively], but, again, sensitivity and specificity were not reported.

### Summary of diagnostic and prognostic accuracy of FibroMAX

No evidence of the diagnostic and prognostic accuracy of FibroMAX was identified.

### Summary of diagnostic and prognostic accuracy of FibroScan

Six studies[117,118,122,127,128,132] assessed the diagnostic and prognostic accuracy of FibroScan in identifying moderate-to-severe fibrosis and cirrhosis in patients with known or suspected ALD ($n = 480$ patients). The study with the most representative population, that by Nahon *et al.*,[128] found that, using a threshold score of 11.6, FibroScan could identify severe (F3–F4) fibrosis with a sensitivity of 87% (95% CI 80% to 93%) and a specificity of 89% (95% CI 76% to 96%), whereas, using a threshold score of 22.7, it could identify cirrhosis with a sensitivity of 84% (95% CI 75% to 91%) and specificity of 83% (95% CI 74% to 91%). As with FibroTest, evidence for FibroScan's

ability to distinguish between patients with and without fibrosis (F1–4 vs F0) is not robust, being based only on one fairly small study by Nguyen-Khac *et al.*[132] ($n = 103$ patients) in a population with a 92% prevalence of fibrosis or cirrhosis.

Two of the included studies, those by Janssens *et al.*[117] and Mueller *et al.*,[118] indicate that FibroScan may overestimate the degree of fibrosis in patients with inflammatory hepatitis. This is consistent with Sagir *et al.*'s[141] finding that 15 out of 20 patients with acute hepatitis of varying causes had FibroScan results indicative of cirrhosis although they had no other signs of cirrhosis, and with Arena *et al.*'s[142] finding that, in patients with chronic hepatitis C, necroinflammatory activity identified at biopsy was associated with increased liver stiffness at each fibrosis stage except cirrhosis. However, unlike Janssens *et al.*,[117] Arena *et al.*[142] found that the degree of steatosis did not influence FibroScan results.

Three studies[95,117,129] ($n \leq 148$ patients) assessed FibroScan's ability to identify PHt. Only one of these, that by Lemoine *et al.*,[97] reported sensitivity and specificity. The study found that FibroScan could identify clinically significant PHt (HPVG $\geq 10$ mmHg) with a sensitivity of 90% (95% CI 78% to 97%) and a specificity of 88% (95% CI 55% to 99%).[97]

A small study by Nguyen-Khac *et al.*[130] ($n = 103$ patients) found that, using a threshold of 47.2 kPa, FibroScan could predict the presence of large oesophageal varices with a sensitivity of 85% (95% CI 67% to 95%) and a specificity of 64% (95% CI 53% to 74%).

### *Discussion of diagnostic and prognostic accuracy of the Enhanced Liver Fibrosis Test, FibroTest and FibroScan*

The evidence relating to the diagnostic accuracy of the ELF test, FibroTest, and FibroScan in relation to liver fibrosis and cirrhosis is not robust, and does not support any attempt to differentiate between their performances in this respect. As Naveau *et al.*[124] note, indirect comparisons between the results of different studies of test accuracy are particularly hazardous, not least because of interstudy variability both in the prevalence of different stages of fibrosis and in biopsy lengths. Only one study was identified that compared two different non-invasive tests with liver biopsy in the same patients: this was the relatively small study by Nguyen-Khac *et al.*,[132] which presented data relating to both FibroTest and FibroScan. In this study, although the point estimates of the AUROCs were higher for FibroScan than for FibroTest, the CIs overlap and, therefore, it is not possible to conclude that FibroScan has better diagnostic accuracy than FibroTest (*Table 27*).

All studies that compare non-invasive tests with liver biopsy in patients with ALD and present information on the interquartile ranges around the median test scores for the different METAVIR stages (i.e. the studies of FibroTest by Nahon *et al.*[128] and Naveau *et al.*,[123] and the studies of FibroScan by Janssens *et al.*,[117] Kim *et al.*,[127] Mueller *et al.*[118] and Nguyen-Khac *et al.*[132]), display a substantial degree of overlap between those interquartile ranges. Thus, for any individual patient, whatever the non-invasive test score, there will be substantial uncertainty regarding their true

**TABLE 27** Comparison of FibroTest and FibroScan with liver biopsy in the same population

| | Condition of interest | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Mild–severe fibrosis (F0–F4) | | Moderate–severe fibrosis (F2–F4) | | Severe fibrosis (F3–F4) | | Cirrhosis (F4) | |
| Test | AUROC | 95% CI | AUROC | 95% CI | AUROC | 95% CI | AUROC | 95% CI |
| FibroTest | 0.77 | 0.63 to 0.90 | 0.79 | 0.69 to 0.90 | 0.80 | 0.70 to 0.91 | 0.84 | 0.72 to 0.97 |
| FibroScan | 0.84 | 0.73 to 0.95 | 0.91 | 0.85 to 0.97 | 0.90 | 0.82 to 0.97 | 0.92 | 0.87 to 0.98 |

fibrosis stage. Although this uncertainty may perhaps be due less to deficiencies in the non-invasive tests themselves than to issues related to liver biopsy (e.g. the use of inadequate samples) or differences between patients in the degree of necroinflammation and steatosis,[143] until well-designed studies are conducted that take these factors into account, the clinical utility of the tests is not apparent.

The evidence relating to the diagnostic accuracy of the ELF test, FibroTest, and FibroScan in relation to PHt and oesophageal varices is weaker than that relating to fibrosis and cirrhosis, as it rests on even smaller patient numbers. Moreover, the use of FibroScan to identify cirrhotic patients at high risk of oesophageal varices is said to be inappropriate because, although varices form only when PHt is present, neither the presence of varices nor their size is directly correlated with the degree of portal pressure elevation.[144]

## Patient management and clinical outcomes

No studies were identified that reported data relating to the effect of the use of any of the four tests on patient management or clinical outcomes.

## Adverse effects and contraindications

The non-invasive tests included in this review appear to be safe. No adverse effects were reported in any of the included studies and no additional evidence has been identified that indicates that transient elastography is specifically associated with any adverse effects. As noted in *The Enhanced Liver Fibrosis Test and FibroTest: failure rates and adverse events*, the ELF test, FibroTest, and FibroMAX, which utilise blood tests, will be associated with the same adverse effects as diagnostic venepuncture generally – primarily pain and bruising, with occasional vasovagal reactions, and very rarely potentially disabling nerve injuries. By contrast, liver biopsy is associated with a high level of morbidity and occasional mortality (see *Chapter 1, Liver biopsy*).

No contraindications have been specified for the ELF test. The contraindications specified for FibroTest, FibroMAX, and FibroScan all relate to the mode of operation of the test, and do not relate to any potential for harm in patients with the relevant characteristics. Moreover, there is evidence to suggest that FibroScan is generally acceptable to patients with ALD. As noted in *FibroScan: failure rates and adverse events*, Janssens *et al.*[117] found that only 2% of patients entering hospital for alcohol detoxification and rehabilitation refused FibroScan, although 34% refused to participate in the study by Nguyen-Khac *et al.*[132] which required them to undergo both FibroScan and blood tests. Finally, in a study of acceptability, Melin *et al.*[145] found that all 380 patients seen for alcohol problems during the course of a year agreed to undergo FibroScan; only 5% (2/44) of those who were offered liver biopsy because their FibroScan result indicated severe fibrosis or cirrhosis refused it, compared with 29% in the study by Janssens *et al.*[117]

## Internal and external validity

The results of the included studies summarised above suggest that the ELF test, FibroTest, and FibroScan can be used to identify patients with ALD who have fibrosis or cirrhosis. However, these results should be viewed with caution for a number of reasons. The most obvious reason is that they are not robust because they rest on data from relatively few patients with ALD; this is especially true of the ELF test.

### *Internal validity*

As noted in *Study quality* above, study quality, as assessed using a modified version of the QUADAS checklist,[112] is generally not high.

Most of the studies display spectrum bias because they recruited patients believed or known to have severe fibrosis or cirrhosis, rather than those representative of the whole spectrum of

patients with suspected ALD. Such spectrum bias favours the index test: because the positive and negative predictive values of diagnostic tests depend critically on the prevalence of the condition being tested for in the population being tested,[146] if the prevalence is considerably higher than would be expected in normal clinical practice, then the positive predictive value of the test will also be higher than it would be in normal clinical practice. Consequently, even if the studies indicate that the tests have high sensitivity and specificity, in normal use many of the positive results will be FPs.[146] Moreover, two of the studies that recruited a more representative patient sample, those by Janssens *et al.*[117] and Melin *et al.*,[122] used the reference standard only in those patients whose index test result was above a specific threshold. This use of the reference standard only in patients testing positive using the index test (verification bias) will result in overestimation of its sensitivity because the number of FN results is too low.[147] In the context of liver fibrosis, both spectrum bias and verification bias are probably due to valid ethical issues surrounding the use of biopsy in patients in whom it is not considered clinically necessary; however, they distort study results in such a way as to favour the index tests.

Conversely, however, studies that compare a non-invasive test with liver biopsy are disadvantaged by the fact that it is an imperfect reference standard; thus, discordance between the degree of fibrosis indicated by biopsy and by non-invasive testing may be because of an error in either test. Mehta *et al.*,[148] noted that liver biopsy is associated with such a degree of potential error that its use as the reference standard may make it impossible to differentiate between a perfect and an inadequate surrogate test. They calculated that, assuming that liver biopsy has a sensitivity and a specificity of 90% for the identification of significant liver fibrosis and that the prevalence of that condition in the population being tested is 40%, a perfect non-invasive test with an AUROC of 0.99 versus true disease can only achieve an AUROC of 0.90 versus liver biopsy. Indeed, Afdhal *et al.*[19] suggest that liver biopsy has a diagnostic accuracy of 80–90% and, in that case, any tests that are compared with liver biopsy cannot achieve an AUROC better than 0.9, and the results are likely to lie in the range 0.75–0.88, with a most likely value of 0.85. Thus, even if a non-invasive test is in fact a perfect non-invasive surrogate for liver biopsy, it may be impossible to prove this.[19]

The use of liver biopsy as the reference standard is associated with a second problem. The non-invasive tests reviewed in this report all present a numeric result relating to a continuous measurement that is held to reflect, directly or indirectly, the degree of fibrosis in the liver. However, this result is then compared with a liver biopsy result expressed in terms of an ordinal scoring system: i.e. biopsy results are classified into a number of groups that have a natural ordering, in that they indicate progressively more severe liver damage, but do not represent a direct arithmetical progression. For example, the degree of fibrosis seen in METAVIR stage F4 is not necessarily twice that seen in METAVIR stage F2; instead, the different stages describe the pattern of deposition of fibrous tissue, as well as its extent.[149] Consequently, to permit comparison with liver biopsy results, a threshold value corresponding to each biopsy stage must be identified for each non-invasive test. In most of the included studies, the threshold values recommended as appropriate for the identification of the different stages of fibrosis and cirrhosis in patients with ALD have been derived statistically from the receiver operating characteristic curve after data collection. They have not been validated prospectively and, therefore, do not fulfil the standard criteria for the general use of a diagnostic test.[138]

### *External validity*

It is difficult to comment on the external validity of the included studies – i.e. the extent to which their populations and methods are generalisable to clinical practice in the UK – not least because of the lack of clarity surrounding the potential role of NILTs in clinical practice in the UK. The issues relate to the population in whom, and the purpose for which, such tests may be used; they are to some extent related.

In the original scope of this assessment, it was envisaged that NILTs would be used in primary care to enable more appropriate selection of patients with abnormal liver function tests and risk factors for chronic liver disease for referral to specialist care. By contrast, the included studies were conducted in secondary or tertiary care settings. Subsequently, clinical experts in the UK have suggested that it is unlikely, and possibly undesirable, that non-invasive tests will be used in primary care, and that most patients who are felt to need further investigation for suspected ALD should be referred to specialist care, where non-invasive tests will be performed if considered appropriate. However, even given this scenario, the range of disease severity is likely to be wider than that seen in the included studies, many of which were limited to patients believed to have relatively severe disease. Indeed, a number of studies recruited patients who not only required liver biopsy for clinical reasons but in whom that biopsy was performed transjugularly rather than percutaneously,[13] suggesting the presence of decompensated cirrhosis.

In ALD, NILTs may be used for one of two main diagnostic purposes:

- to identify patients with fibrosis, so that efforts may be made to prevent the development of cirrhosis
- to identify patients with cirrhosis, enabling them to be monitored for the development of conditions such as oesophageal varices and HCC.

Assuming that non-invasive test results indicative of fibrosis are effective in influencing patients with ALD to abstain from alcohol – and no evidence for this has been identified – then the former use is of potentially greater clinical value as it would permit the identification of patients with ALD at a time when that disease was still reversible, whereas identification of patients with cirrhosis would only permit the initiation of monitoring to enable prompt treatment of symptoms of an incurable disease. However, only one study, that by Nguyen-Khac *et al.*,[132] reported on the ability of a non-invasive test to identify mild (METAVIR F1) as well as more severe (F2–F4) fibrosis in patients with suspected ALD, and in most studies the tests performed better when identifying cirrhosis (F4) than when identifying mild (F1), moderate (F2), or severe (F3) fibrosis. This clearly limits the clinical utility of the tests.

In tests that present results derived from a continuous scale, the intended purpose of that test will affect the choice of the threshold score. So, if the intended purpose of the NILTs reviewed in this report is to identify patients with cirrhosis to undergo further tests and monitoring, a threshold score should be chosen that maximises sensitivity (i.e. the proportion of patients who genuinely have the condition of interest who are correctly identified by the non-invasive test), as this will minimise the risk of patients with cirrhosis being mistakenly identified as not having the condition, and therefore not receiving further tests, monitoring, and treatment, as appropriate. However, if the intended purpose of the tests is to exclude patients without fibrosis, the threshold score should be chosen to maximise specificity (i.e. the proportion of people who genuinely do not have the condition of interest who are correctly identified by the non-invasive test), to reduce the risk of patients who do not have fibrosis undergoing costly and potentially invasive tests.

Test results may be influenced by factors other than the degree of fibrosis present in the liver. The included studies have shown that, for FibroScan, the optimum threshold values for fibrosis and cirrhosis are higher in patients with ALD than in patients with hepatitis C (and possibly other liver diseases), and it is therefore crucial that the aetiology of suspected liver disease is securely established before the test result is interpreted.[18] In addition, as noted in *FibroScan: diagnostic accuracy results*, in patients with a secure diagnosis of ALD, FibroScan may overestimate the degree of fibrosis if either steatosis or alcoholic hepatitis is present. Current drinking status is also relevant: Mueller *et al.*[118] have shown that, in patients with ALD, liver stiffness, as measured

by FibroScan, decreases during alcohol detoxification independent of the fibrosis stage. Thus, consideration must also be given to the optimum timing of the tests.

Finally, it should be noted that, unlike liver biopsy, the non-invasive tests assessed in this report only seek to identify the degree of liver fibrosis. They cannot also provide additional useful information, for example by indicating the presence of another liver disease in addition to ALD, or by evaluating necroinflammation to assess whether that fibrosis is an ongoing process that may continue to develop or whether it results from a past event that has stabilised or even regressed.[104]

### *Conclusions for clinical effectiveness*

There is some evidence to suggest that, in patients with known or suspected ALD, the ELF test, FibroTest, and FibroScan can identify fibrosis with varying degrees of diagnostic accuracy; no evidence has been identified relating to FibroMAX, although this is recommended by the manufacturers in preference to FibroTest in patients with ALD. Although FibroTest and FibroScan appear to have greater accuracy in identifying cirrhosis rather than lesser degrees of fibrosis, the ELF test appears to perform less well in specifically identifying cirrhosis than in identifying the presence of moderate-to-severe fibrosis but, as the evidence base is very small and acceptable minimum standards were not used for the biopsy samples, this finding is not robust. Evidence for the ability of FibroTest and FibroScan to identify clinically significant PHt, and oesophageal varices, rests on extremely small studies, and again is not robust.

Moreover, the confidence that can be placed in the study results is reduced because most of the studies display spectrum bias, and the two studies that recruited a more representative sample display verification bias: both of these biases will favour the index test. In addition, the degree of error associated with liver biopsy is such that its use as the reference standard may make it impossible to judge with accuracy the adequacy of the surrogate test. Finally, the degree of overlap between the interquartile ranges around the median values relating to each METAVIR stage means that, for any individual patient, whatever their non-invasive test score, there will be substantial uncertainty regarding their true fibrosis stage, and this will substantially limit the clinical utility of the non-invasive tests.

# Chapter 5

# Cost-effectiveness: model parameters

This chapter details the parameters within the mathematical model and the sources used to provide the values assumed in the analyses. There was a considerable number of data that were not available and broad assumptions have been made to allow an estimation of the range of the potential cost-effectiveness of each NILT after considering the management and likely life expectancy of a patient following diagnosis. Sensitivity analyses have been performed in order to test the robustness of the results produced to changes in the input parameters.

## Discount rates

In accordance with the NICE methods guide,[150] both benefits and costs have been discounted at a rate of 3.5% per annum.

## The sensitivity and specificity of non-invasive liver tests used within the economic model

As detailed, the sensitivity of each NILT, alongside that of biopsy, will determine the number of patients who have cirrhosis who are not appropriately diagnosed. The specificity of the NILT will determine the number of patients who may receive unnecessary biopsy or, if a triaging strategy is not pursued, the number of patients who are monitored unnecessarily. The advantage of a NILT is that fewer biopsies will be performed than in assumed current practice, which will be associated with reduced costs, and also reduced mortality and morbidity.

The estimated sensitivity and specificity for each NILT have been described previously in *Tables 15*, *16* and *19*. No formal meta-analysis has been undertaken because of the potential heterogeneity within the trials in terms of the length of and the number of portal tracts examined within the liver biopsy, the number of days between the test and performing the biopsy, the level of current drinking within the cohort and the different cut-off thresholds for diagnosing cirrhosis/fibrosis. In order to provide an estimation of the cost-effectiveness of the NILTs, three scenarios for sensitivity and specificity are evaluated, which the authors have selected from the combinations of RCTs and cut-off thresholds for classification of cirrhosis.

These results will be indicative of the likely cost-effectiveness, although some caveats must be provided. These include:

1. The fact that the sensitivity and specificities reported are directly calculated from the trial data, whereas ideally the results would be calculated from a threshold that was specified in advance of the trial.
2. Inconsistency between reported trial data, for example, considering FibroTest, the study by Naveau *et al.*[124] has both a higher sensitivity and specificity at a 0.70 threshold than the study by Thabut *et al.*, 2007a[13] which used a threshold of 0.74.
3. Small patient numbers and their fibrosis levels mean that there is reasonably large uncertainty. Where values for TPs, FNs, TNs and FPs were < 5, a non-informative prior of

0.5 was added to that value and the corresponding value when calculating sensitivity and specificity, which is equivalent to using Jeffreys' prior. As such, the values for sensitivity and specificity may not match exactly those reported in the systematic review section.
4. The fact that the ELF did not report sensitivity and specificity for detecting only cirrhosis and that these have been inferred from a moderate/severe fibrosis population.

In addition to these caveats, there is the possibility that biopsy may not be a perfect gold standard and that sampling error may result in a FN being indicated by the test. As previously reported, data obtained when comparing liver biopsy with post-mortem examination showed that cirrhosis was detected in only 16 out of 20 cases when a single biopsy was taken, although this increased to 20 out of 20 when three biopsies were taken.[91] Further data show that when three biopsies were taken using the same entry site the sensitivity was 50%.[92] Exploratory results have been undertaken assuming that the sensitivity of biopsy in detecting cirrhosis is 80% rather than 100%. In this instance, further assumptions need to be made as the NILTs were compared with biopsy. Three scenarios have been explored, in a similar manner as previously used by Carlson *et al.*[151]

1. That the NILTs also failed to detect the cirrhoses missed by the biopsy. This is termed a pessimistic scenario.
2. That the NILTs detected the cirrhoses missed by the biopsy, but these were recorded in the study as FPs. This is termed an optimistic scenario.
3. That the NILTs also failed to detect a proportion of the cirrhosis missed by the biopsy, with the remaining proportion being recorded as FPs. The proportion also missed would be assumed to be 1 – sensitivity reported in the trial. This is termed a stochastic scenario.

An example of this methodology is given below using data for detecting cirrhosis by FibroTest as reported by Naveau *et al.*[123] This trial reported (compared with a gold standard of biopsy) 62 TPs, 6 FNs, 133 TNs and 20 FPs, with a sensitivity of 91%. Biopsy detected 68 positives, but if the sensitivity of biopsy was 80% then biopsy would be expected to miss 17 (68/4) actual positives.

In the pessimistic scenario it would be assumed that 17 cases of TNs were actually FNs. Thus, the true distribution was 62 TPs, 23 FNs, 116 TNs and 20 FPs.

In the optimistic scenario it would be assumed that 17 cases of FNs were actually TPs. Thus, the true distribution was 79 TPs, 6 FNs, 133 TNs and 3 FPs.

In the stochastic scenario, it would be assumed that 91% of the 17 cases missed by biopsy would be detected and were classed as FPs; this number is rounded to the nearest integer and down if equidistant. This would result in 15 being initially recorded as FPs and 2 as TNs. Thus, the true distribution was 77 TPs, 8 FNs, 131 TNs and 5 FPs.

This approach was used for all test scenarios. In circumstances where the results could not meet the decision rules, for example if there were only three FNs but it was expected that biopsy would miss 10 that would all be diagnosed correctly in the optimistic scenario, then the maximum number that could be transferred from FPs to TPs would be transferred, which would be three in this example. When calculating the resultant sensitivities and specificities, if values for TPs, FNs, TNs and FPs were below 5, an uninformative prior of 0.5 was added to that value and the corresponding value when calculating sensitivity and specificity.

The sensitivities and specificities used in the model for each scenario are provided in *Tables 28–31*. It is noted that where there are few patients that biopsy rated as without cirrhosis, for instance

**TABLE 28** Test characteristics for detecting cirrhosis used within the model when it is assumed that biopsy has 100% sensitivity and 100% specificity

| NILT | Scenario | Calculated from | Threshold | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|
| ELF | 1 | Rosenberg 2004[121] | 0.431 | 92.4 | 97.4 |
| | 2 | Rosenberg 2004[121] | 0.087 | 98.9 | 18.4 |
| | 3 | Authors' estimate correcting for a cirrhotic population only | 0.431 | 96.0 | 90.0 |
| FibroTest | 1 | Naveau 2005[123] | 0.70 | 91.2 | 86.9 |
| | 2 | Naveau 2005[123] | 0.30 | 99.3 | 50.3 |
| | 3 | Thabut 2007[13] | 0.74 | 78.3 | 78.6 |
| FibroScan | 1 | Mueller 2010[118] | 11.5 | 98.1 | 77.3 |
| | 2 | Janssens 2010[117] | 19.6 | 78.6 | 75.9 |
| | 3 | Nahon 2008[128] | 22.7 | 83.5 | 83.6 |
| Biopsy | – | Assumption | N/A | 100.0 | 100.0 |

N/A, not applicable.

**TABLE 29** Test characteristics for detecting cirrhosis used within the model when it is assumed that biopsy has 80% sensitivity and 100% specificity and a pessimistic scenario is employed

| NILT | Scenario | Calculated from | Threshold | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|
| ELF | 1 | Rosenberg 2004[121] | 0.431 | 75.0 | 93.8 |
| | 2 | Rosenberg 2004[121] | 0.087 | 92.9 | 3.1 |
| | 3 | Authors' estimate correcting for a cirrhotic population only | 0.431 | 78.0 | 87.0 |
| FibroTest | 1 | Naveau 2005[123] | 0.70 | 72.9 | 85.3 |
| | 2 | Naveau 2005[123] | 0.30 | 80.0 | 44.1 |
| | 3 | Thabut 2007[13] | 0.74 | 78.3 | 25.0 |
| FibroScan | 1 | Mueller 2010[118] | 11.5 | 81.3 | 75.3 |
| | 2 | Janssens 2010[117] | 19.6 | 64.0 | 70.8 |
| | 3 | Nahon 2008[128] | 22.7 | 66.7 | 76.6 |
| Biopsy | – | Assumption | N/A | 80.0 | 100.0 |

N/A, not applicable.

in Thabut *et al.* 2007a,[13] where only six patients were diagnosed as not cirrhotic by biopsy, the fluctuations in test characteristics can be large, and that in the pessimistic scenarios the sensitivities of the NILTs may become better than that for biopsy.

For all scenarios, the sensitivity of clinical experience alone was 81.3% and specificity was 89.2%.[9] For diagnosing all patients with cirrhosis sensitivity was 100% and specificity 0%.

Examination of *Table 31,* where a stochastic scenario has been assumed, shows that this is often identical or close to those predicted in the optimistic scenario (see *Table 30*), which produces results that are more favourable to the tests. Given this, for brevity reasons it was decided that results would not be produced using the stochastic scenario, with the authors comfortable that the remaining scenarios (100% sensitivity for biopsy, pessimistic and optimistic) provided a good indication of the uncertainty within the decision.

**TABLE 30** Test characteristics for detecting cirrhosis used within the model when it is assumed that biopsy has 80% sensitivity and 100% specificity and an optimistic scenario is employed

| NILT | Scenario | Calculated from | Threshold | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|
| ELF | 1 | Rosenberg 2004[121] | 0.431 | 92.4 | 97.4 |
| | 2 | Rosenberg 2004[121] | 0.087 | 99.1 | 43.8 |
| | 3 | Authors' estimate correcting for a cirrhotic population only | 0.431 | 95.0 | 92.0 |
| FibroTest | 1 | Naveau 2005[123] | 0.70 | 92.9 | 97.4 |
| | 2 | Naveau 2005[123] | 0.30 | 99.4 | 56.6 |
| | 3 | Thabut 2007[13] | 0.74 | 78.7 | 91.7 |
| FibroScan | 1 | Mueller 2010[118] | 11.5 | 98.5 | 84.1 |
| | 2 | Janssens 2010[117] | 19.6 | 90.0 | 82.7 |
| | 3 | Nahon 2008[128] | 22.7 | 85.6 | 99.1 |
| Biopsy | – | Assumption | N/A | 80.0 | 100.0 |

N/A, not applicable.

**TABLE 31** Test characteristics for detecting cirrhosis used within the model when it is assumed that biopsy has 80% sensitivity and 100% specificity and a stochastic scenario is employed

| NILT | Scenario | Calculated from | Threshold | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|---|
| ELF | 1 | Rosenberg 2004[121] | 0.431 | 90.4 | 97.2 |
| | 2 | Rosenberg 2004[121] | 0.087 | 99.1 | 43.8 |
| | 3 | Authors' estimate correcting for a cirrhotic population only | 0.431 | 94.0 | 90.0 |
| FibroTest | 1 | Naveau 2005[123] | 0.70 | 90.6 | 96.3 |
| | 2 | Naveau 2005[123] | 0.30 | 99.4 | 56.6 |
| | 3 | Thabut 2007[13] | 0.74 | 75.0 | 84.3 |
| FibroScan | 1 | Mueller 2010[118] | 11.5 | 98.5 | 84.1 |
| | 2 | Janssens 2010[117] | 19.6 | 80.0 | 86.0 |
| | 3 | Nahon 2008[128] | 22.7 | 82.8 | 99.1 |
| Biopsy | – | Assumption | N/A | 80.0 | 100.0 |

N/A, not applicable.

## The prevalence of cirrhosis in the defined population

In addition to the test characteristics, the prevalence of cirrhosis in people whom a secondary care clinician would want to biopsy is needed to determine the absolute number of TPs, FPs, TNs and FNs. Based on clinical advice this value has been set to 35%.

## The costs of biopsy and each non-invasive liver test

The costs used within the model are provided in *Table 32*. These have been inflated, where applicable, using the inflation indices reported in Curtis.[152] These costs are deemed additional to standard clinical practice.

Sensitivity analyses were undertaken to examine the robustness of the results to changes in the prices of diagnostic tests.

**TABLE 32** The estimated cost of each test (2008–9 prices)

| Test | Cost (£) | Source |
|---|---|---|
| Percutaneous biopsy | 894 | Stamuli 2009[56] |
| Transjugular biopsy | 1500 | An indicative figure for a biopsy requiring an overnight stay and possible transportation costs[9] |
| ELF | 45 | Clinical input. This value comes from an early adopter quote provider to the Royal Free Hospital, London, UK for 100 ELFs per month. (Dr Marsha Morgan, Royal Free Hospital, London, UK, 2010, personal communication) |
| FibroScan (marginal cost) | 50 | Clinical input suggests that this is likely to be the price charged to the NHS per scan. This is preferred to the Stamuli 2009[56] estimated cost of £19.52 (range £12.44–33.94) |
| FibroTest | 50 | Set similar to the cost of the ELF as both are blood tests and are likely to be competitively priced. This estimate is preferred to a value of €90–300 reported by Morra 2007[67] |
| Clinical experience | 0 | Assumption |
| Diagnosing all with cirrhosis | 0 | Assumption |

## Adverse events related to each diagnostic test

It has been assumed that no NILT has adverse events, aside from a potential misdiagnosis of cirrhosis. Biopsy, however, owing to its invasive nature, is associated with both mortality and morbidity. From the systematic review undertaken in this study, we have assumed that percutaneous biopsy has a probability of 0.09% of causing mortality, with an additional risk of causing a serious adverse event of 0.72% (see *Appendix 3*). The corresponding values for transjugular biopsy are 0.18% and 1.27%, respectively (see *Appendix 3*).

A serious adverse event was deemed to be associated with a hospital stay, assumed to cost £1000, and a QALY decrement of 0.2 (equivalent to approximately 10 weeks with a utility of zero or a year with a utility decrement of 0.2). The QALY value was arbitrary, but was assumed to be a value that would be likely to disfavour biopsy; sensitivity analyses were undertaken to assess the robustness of the results to changes in both the costs and QALY decrements assumed to be related to serious adverse events.

Applying the base-case values to the risk of a serious adverse event results in an expected QALY decrement per patient of 0.000142 for a percutaneous biopsy and 0.000254 for a transjugular biopsy; the cost implications per patient would be £7 for a percutaneous biopsy and £13 for a transjugular biopsy. For all patients who die as a result of biopsy, the costs of the biopsy are assumed to be incurred, but no further QALYs will be accrued. It is uncertain whether or not patients undergoing a biopsy will suffer anxiety prior to the procedure; in order to address this issue and to assess the robustness of the results to this assumption, a sensitivity analysis was performed where the disutility associated with a biopsy was increased to 0.04 QALY (a value equivalent to approximately a fortnight with a utility of zero), which was deemed in consultations with clinicians to be an upper bound.

## The proportion of tests that will produce results that cannot be used

The results of the NILTs can be confounded by patient characteristics such as obesity and concurrent drinking. This has been reflected within the model by assuming that the rates of tests that cannot be used are 20% for FibroScan and 25% for both blood tests (FibroTest and the ELF). The model assumes that when a test has produced a result that cannot be used, the patient will then receive a biopsy.

## The outcomes associated with each final node within the economic model for diagnosis of cirrhosis

*Figures 1–3* detail the assumed pathways within the model. These strategies have four common end points: (1) abstinent following a diagnosis of cirrhosis; (2) continuing to drink following a diagnosis of cirrhosis; (3) abstinent following a diagnosis of no cirrhosis; and (4) continuing to drink following a diagnosis of no cirrhosis. Those strategies that incorporate a biopsy also have the risk of biopsy mortality. Excluding biopsy mortality, these common end points represent an amalgamation of heterogeneous patient experiences into a long-term estimation of costs and QALYs. These end points are also further broken down into whether or not the diagnosis received was correct, which will differ depending on whether or not the diagnosis regarding cirrhosis is correct, increasing the actual number of end points within the model to eight.

Although not depicted within the model diagram, a proportion of patients who do not have cirrhosis at the time of the investigation but who continue to drink heavily are assumed to develop cirrhosis. These patients will be assigned the costs and QALYs associated with cirrhotic patients who continue to drink heavily rather than non-cirrhotic patients who continue to drink heavily.

Because of the likely heterogeneity of patients within each of the eight end points, the reliability of any estimate will be questionable. To populate the model we have collated data from various sources, including clinical input, and provided relatively broad estimates to guard against spurious accuracy. The sensitivity of the modelled results to these values is tested within sensitivity analyses. Each end point is discussed individually and, for reference, the cost and QALY values for each of the end points are collated in *Table 33*.

### *End point 1: true-positives for cirrhosis who abstain from drinking*

These patients will receive monitoring for HCC, HE and oesophageal varices. The estimated costs and QALYs associated with patients who are screened for HCC have been taken from a study by Thompson Coon *et al.*,[153] which reports that ALD patients with cirrhosis undergoing annual serum α-fetoprotein and 6-monthly ultrasound scans were estimated to accrue 9.410 QALYs at a cost of £27,400. This source was selected as it came from a health technology assessment that had explicitly divided the cost of surveillance for HCC into aetiology types, allowing values for ALD patients to be explicitly used.

**TABLE 33** The lifetime costs and QALYs associated with each end point including the costs of electroencephalograms and the costs and QALY implications of screening for varices, providing prophylaxis treatment where appropriate and treating variceal bleeding

| End point | Costs (£) | QALYs |
|---|---|---|
| TPs for cirrhosis who abstain from drinking | 29,980 | 9.679 |
| TPs for cirrhosis who continue drinking | 39,474 | 4.399 |
| FPs for cirrhosis who abstain from drinking | 25,154 | 11.066 |
| FPs for cirrhosis who continue drinking[a] | 25,154 | 11.066 |
| TNs for cirrhosis who abstain from drinking | 1000 | 11.066 |
| TNs for cirrhosis who continue drinking[a] | 1000 | 11.066 |
| FNs for cirrhosis who abstain from drinking | 26,100 | 9.359 |
| FNs for cirrhosis who continue drinking | 36,100 | 3.744 |

a A proportion of these patients are assumed to progress to cirrhosis and incur the costs and QALYs of TPs who continue to drink.
See *Chapter 5, The outcomes associated with each final node within the economic model for diagnosis of cirrhosis* for an explanation of the derivation of the values within the table.

### End point 2: true-positives for cirrhosis who continue drinking

Clinical advice indicates that such patients will still receive monitoring for HCC, HE and varices despite their non-abstinence. Clinical advice also indicates that the costs associated with such patients will be much higher as the risk of progressing to more severe states, such as decompensated cirrhosis, is greater in the non-abstinent, as is the risk of HE and ascites. The literature has suggested that the cost per annum of decompensated cirrhosis is £9385 compared with £1171 per annum for compensated cirrhosis.[154] The costs of HE are £2718 per event (Code AA22Z),[155] and the treatment of ascites has been reported to cost US$4048 per annum.[156] All historic costs have been inflated to 2008–9 costs.[152] The sources of these costs were selected as these were also used in the Thompson Coon *et al.* study[153] to provide consistency. It was assumed that these costs would approximate to an additional £10,000 per person when compared with TPs who were abstinence and a cost of £37,400 was used.

For QALYs accrued, a study[21] reported that the survival rate at 7 years for patients with cirrhosis was 72% for abstainers and 44% for those who continued to drink. For simplicity, it was assumed that these were results from exponential distributions; this would indicate that cirrhotic patients who continued to drink would have (on average) only 40% of the life expectancy of cirrhotic patients who abstained. Assuming that this value could also be used to downgrade QALYs, it is estimated that the QALYs accrued by TPs who continued to drink would be 3.764 ($9.410 \times 40\%$).

### End point 3: false-positives for cirrhosis who abstain from drinking

These patients will receive monitoring for HCC, HE and oesophageal varices. The costs for these patients have been estimated from the costs of TPs who abstain (£27,400), but with a consideration that FPs are likely to incur fewer costs because of avoiding cirrhosis. It is assumed that FPs would be representative of the group termed by Wright *et al.* as having moderate disease,[154] which was defined as having either a fibrosis score of between 3 and 5 or a necroinflammatory score of $> 3$. The cost per annum was £421 less in the group with moderate disease than in those with cirrhosis, a difference which has been increased to £469 per annum to incorporate inflation. Data from Wright *et al.*[154] were selected, as these were also used in the Thompson Coon *et al.* study[153] and provided some consistency.

It is assumed that these costs would be applied for 20 years, resulting in a discounted cost reduction of £6899 and an overall cost of approximately £20,500. In the absence of further data, we have assumed that the QALYs gained will be equivalent to that of TNs, 11.066, using the methodology described later (see *End point 5: true-negatives for cirrhosis who abstain from drinking*).

### End point 4: false-positives for cirrhosis who continue drinking

Few data are available for FPs who continue to drink. We have conservatively assumed that, for the majority of patients, the costs are equivalent for FPs who abstain (£20,500), as are the QALYs accrued (11.066). However, to acknowledge the fact that such patients may develop cirrhosis, a proportion are modelled to have the cost and QALY implications associated with TPs who continue to drink rather than those for FPs.

### End point 5: true-negatives for cirrhosis who abstain from drinking

These patients will receive lifestyle advice only. A cost of £1000 per patient was estimated (based on clinical advice), to account for appropriate monitoring of the patient. In order to estimate the QALYs accrued we have assumed that the patients are the same age as those in Thompson Coon *et al.*[153] (53 years) and that they will have a reduced life expectancy compared with general population, but will live for a further 20 years. Using European Quality of Life-5 Dimensions (EQ-5D) population norms[157] and assuming discounting at 3.5% per annum,[150] an estimate of 11.066 QALYs was derived.

### *End point 6: true-negatives for cirrhosis who continue drinking*

These patients will receive lifestyle advice only. We have assumed that the costs and QALYs accrued are the same as for TNs who are abstinent (£1000 and 11.066, respectively). It has been assumed that such patients may develop cirrhosis; thus, a proportion are modelled to have the cost and QALY implications associated with TPs who continue to drink rather than those for TNs.

### *Endpoint 7: false-negatives for cirrhosis who abstain from drinking*

These patients will receive lifestyle advice only. The estimated costs and QALYs associated with ALD patients with cirrhosis who are not screened for HCC have been taken from a study by Thompson Coon *et al.*,[153] which reports that patients were estimated to accrue 9.359 QALYs at a cost of £26,100.

### *End point 8: false-negatives for cirrhosis who continue drinking*

These patients will receive lifestyle advice only. It is uncertain how much additional cost would be incurred by this group compared with FNs who abstain; however, following the logic described for TPs, a value of £10,000 was assumed, resulting in a cost of £36,100. The QALYs accrued have been set at 40% of the level of FNs who abstain, using the reasoning described in the 'TPs who continue drinking' section. This would equate to 3.744 QALYs (9.359 × 40%).

## The proportion of patients without cirrhosis who continue to drink heavily that will develop cirrhosis

On clinical advice this has been set to 20%.

### *Incorporating oesophageal varices bleeding*

The analyses of screening for HCC undertaken by Thompson Coon *et al.*[153] did not include any costs or consequences associated with oesophageal varices bleeding. However, these can have significant morbidity and can also cause mortality. This has been included within the modelling by estimating the effects and costs associated with screening for varices, providing prophylaxis treatment where appropriate, and with bleeds in a separate model and then including these within our analyses. The varices model used the following assumptions, which were based, where possible, on recommendations provided by de Franchis.[158]

1. All patients with cirrhosis would receive an endoscopy at a cost of £200 (data from the Royal Free Hospital, London, UK).
2. Forty per cent of those with cirrhosis will have varices,[43] of which one-third are small, one-third moderate and one-third large (clinical advice).
3. The risk of bleeding in those who continue to drink is high, with a risk of 15% per annum,[156,159] although this risk is reduced by 50% if prophylactic treatment is provided.[159]
4. On clinical advice, the risk of bleeding among those who abstain was reduced to 5%, with a reduced risk of 2.5% in those taking prophylactic treatment.
5. Only those patients who have been diagnosed (correctly or incorrectly) with cirrhosis will be provided with prophylactic treatment. Patients with cirrhosis who were not correctly diagnosed will be provided with prophylactic treatment following an oesophageal varices bleed.
6. Prophylactic treatment with propranolol 40 mg twice daily costing £28 per year[160] is provided to all patients with varices, excluding 50% of those with small varices (clinical advice).

7. The risk of mortality following a bleed is assumed to be 30%[43] and unaffected by prophylactic treatment.
8. The cost of treating an oesophageal varices bleed is assumed to be £10,000. This has been estimated from a weighted average of costs presented by Thabut *et al.*,[161] which was €11,982; these costs are similar to figures reported in Wechowski *et al.*[162]
9. An endoscopy once every 3 years to be provided for those without varices and every 2 years for those patients with small or moderate varices.
10. Those patients with large varices would have three sessions of band ligation costing £438 (US$674) each, with an endoscopy at 6 months and then at yearly intervals.[162]
11. Patients falsely diagnosed as not having cirrhosis who have a subsequent variceal bleed would be prescribed propranolol 40 mg twice daily.

Compared with a no screening strategy, these assumptions provide an estimated net incremental cost of screening and subsequent treating to a patient with oesophageal varices who continues to drink of £400 and of £900 for those who abstain. The additional QALYs accrued from screening and providing subsequent treatment compared with a no screening strategy were 0.635 for those patients who continue to drink and 0.269 for abstinent patients. For simplicity within the model, these cost and QALY values were added to end point 1, for those people who abstain from alcohol, and end point 2 for those people who continue to drink heavily.

### Detecting hepatic encephalopathy

An electroencephalogram [estimated to cost £220 (data obtained from the Royal Free Hospital, London, UK)] was assumed to be required annually to detect HE. These were estimated to have a discounted cost of £1674 (US$2578) over a lifetime for those patients diagnosed with cirrhosis.

### The proportion of patients who continue to drink in relation to diagnosis by biopsy

One small study ($n = 96$), published in abstract form,[26] reported that the level of continued drinking following biopsy-proven ALD was dependent on whether or not the patient had cirrhosis. Fifty-nine per cent of patients without cirrhosis had a heavy intake compared with 38% of those who had cirrhosis, a difference which was significant ($p = 0.04$). As it is plausible that the abstinence rate may be higher in patients with cirrhosis, these values were used within the model.

### The proportion of patients who continue to drink in relation to diagnosis by a non-invasive liver test

It is conceivable that patients who have a less invasive test may be more reluctant than those who have a biopsy to become abstinent. Possible reasons for this include (1) the knowledge of the physician (and potentially a well-informed patient) that the NILT has low specificity; (2) the fact that inflammation and fat can be assessed directly following a biopsy (but not through a NILT), which can be used to inform the patient of an underlying disease progression that has not currently reached cirrhosis or severe fibrosis level; and (3) potentially the fact that a biopsy is invasive may also convince a patient that the disease is potentially life-threatening.

Exploratory results showed that, unsurprisingly, the model was very sensitive to the level of abstinence achieved in the ALD patients. In order to provide meaningful results, the abstinence rates were subjected to a threshold sensitivity analyses, where the percentage point increase in abstinence rates, for both those diagnosed with cirrhosis and those diagnosed without cirrhosis, required to change a decision regarding the cost-effectiveness of a strategy was calculated.

## The benefit of biopsy in identifying liver disease that is not alcohol-related liver disease and confirming alcohol-related liver disease

Biopsy allows the aetiology of the liver disease to be established; this is not possible when NILTs alone are used. A diagnosis of liver disease that is not ALD would allow patients to receive alternative treatment for this condition rather than for ALD. It may also be the case that a patient may derive some underlying benefit from a definitive diagnosis of his or her condition. Initial exploratory results indicated that the results of the model were extremely sensitive to any gains in QALYs owing to a biopsy being performed. As such, this value has been left at zero for the analyses, but is subject to threshold sensitivity to determine the value that this would need to be greater than in order for the decision on which treatment was the most cost-effective to change.

# Chapter 6

# Cost-effectiveness model: results

During the course of the project, exploratory results were produced that showed that the model was relatively insensitive to some parameters, but was very sensitive to some parameters on which there were few data.

Those parameters that were shown to have little influence on the modelled results were subjected to univariate sensitivity analyses using a set of pre-defined ranges to check the robustness of the results regarding the cost-effectiveness of each. These results and the ranges used are shown in *Table 34*. These results assumed that the abstinence rates were independent of the diagnostic test and that a biopsy conferred no QALY gain and used scenario 1 (described in *Table 35*). In this scenario, only a FibroScan triage policy was more cost-effective than biopsying all patients, assuming a cost per QALY threshold of £20,000, where the cost-effectiveness of biopsying all patients had a cost per QALY of £71,327 compared with a FibroScan triage policy. The cost per QALY of biopsying all patients was £15,683 compared with an ELF triage policy and £17,702

**TABLE 34** The univariate sensitivity analyses conducted using scenario 1, and assuming equal rates of abstinence in those diagnosed with cirrhosis and assuming no incidental benefit

| Parameter | Central estimate | Extreme values tested | Cost-effectiveness[a] |
|---|---|---|---|
| Prevalence of cirrhosis (%) | 35 | 25–45 | When set to 25%, the ELF triage strategy (£26,547) and the ELF replacement strategy (£28,783) became cost-effective |
| Cost of a percutaneous biopsy (£) | 894 | 600–1500 | When set to £1500, the ELF triage strategy (£26,581) and the ELF replacement strategy (£38,245) became cost-effective |
| Costs of FibroScan, ELF and FibroTest (£) | 50, 45 and 50 | 40–100 | No change |
| Simultaneous changing of all costs associated with the eight end points (%) | See *Table 33* | –25–50 | When set to –25%, the ELF replacement strategy (£20,404) became cost-effective |
| Simultaneous changing of all QALYs associated with the eight end points (%) | See *Table 33* | –25–50 | When set to –25%, the ELF triage strategy (£20,928) and the ELF replacement strategy (£23,641) became cost-effective |
| Mortality rate following percutaneous biopsy (%) | 0.09 | 0.05–0.20 | When set to 0.2%, the ELF replacement strategy (£27,293) became cost-effective |
| The average costs of dealing with complications following a biopsy (£) | 7 | 0–100 | No change |
| Percentage of biopsies that need to be repeated (%) | 0 | 0–5 | No change |
| Percentage of FibroScans that do not produce usable results (%) | 20 | 10–40 | No change |
| Percentage of FibroTests that do not produce usable results (%) | 25 | 10–40 | No change |
| Percentage of ELFs that do not produce usable results (%) | 25 | 10–40 | No change |
| The proportion of patients who progress to cirrhosis if they continue to drink heavily (%) | 20 | 10–40 | When set to 40%, the ELF replacement strategy (£23,440) became cost-effective |

a Changes to the cost-effectiveness of strategies compared with biopsying all patients assuming a willingness to pay of £20,000 per QALY. The numbers in parentheses indicate the cost per QALY of biopsying all compared with the strategy, where the cost-effectiveness conclusion has changed.

compared with an ELF replacement policy, which were the two strategies where the cost-effectiveness of biopsying all patients was most uncertain.

In a number of univariate sensitivity analyses, there was a change in the cost-effectiveness of some strategies compared with biopsying all patients, although often this was because of the cost-effectiveness of biopsying all patients becoming slightly > £20,000 per QALY. Given the large impact on results caused by plausible changes in the remaining parameters, a decision was made to simplify the results produced by fixing the parameters in *Table 34* at their central estimates, with the acknowledgement that uncertainty and interactions between these parameters were not considered, which would result in the results produced underestimating uncertainty.

Strategies with poor specificity for cirrhosis (diagnosing all with cirrhosis, and in some scenarios the ELF) produced the greatest number of QALYs owing to the assumed higher proportion of patients who became abstinent because of a diagnosis of cirrhosis. However, these relatively small QALY gains were not sufficient to be deemed cost-effective when the costs of follow-up in patients incorrectly diagnosed with cirrhosis were considered.

Key parameters that were identified as affecting the cost-effectiveness were:

1. The assumed sensitivity of biopsy and the assumption made regarding the accuracy of the NILTs in detecting these FPs. The scenarios analysed were 100% sensitivity, 80% sensitivity with pessimistic assumptions for the NILTs and 80% sensitivity with optimistic assumptions for the NILTs.
2. The assumed sensitivity and specificity of biopsy and of each NILT. These were the three scenarios detailed in *Tables 28–30*.
3. Whether or not the biopsy was undertaken using a percutaneous or transjugular method.
4. Whether or not the QALY loss associated with a biopsy would include any effects of anxiety. This was denoted as normal, which was the estimated value associated with severe adverse events, or high, a value of 0.04.
5. The assumed changes in abstinence rates dependent on the tests used to diagnose cirrhosis.
6. The assumed potential incidental QALY benefits of biopsy.

The fifth and sixth parameters within the list were evaluated using a threshold approach for each combination of the remaining four parameters. This resulted in 36 potential scenarios, which are detailed in *Table 35*. It is stressed that the scenarios regarding sensitivity and specificity for each of NILTs do not have to be compared directly, and that it is possible to compare scenario 1 for one NILT with scenario 2 for another, and with scenario 3 for a third. Similarly, when 80% sensitivity is assumed for biopsy, it may be plausible that one NILT could use an optimistic assumption, whereas a separate test, with a different modality, may use a pessimistic assumption.

Given the large uncertainty in the two parameters shown to be key in the exploratory analyses (the relationship between abstinence and the test used to diagnose cirrhosis and the gain in QALYs associated with biopsy), it was decided that carrying out a formal probabilistic sensitivity analyses[163] would offer little additional insight.

The results in terms of costs, QALYs and predicted number of patients who receive a biopsy are presented in *Appendix 10* for each of the 36 scenarios. These tables can be used to calculate incremental cost-effectiveness ratios for comparisons of all scenarios presented; the range of cost-effectiveness values is wide and can be seen to be favourable to the use of NILTs in some scenarios, with the use of NILTs producing more QALYs at a lower cost, but favourable to biopsy in other scenarios. However, it is stressed that these values assume that there is no decrease in abstinence rates when NILTs are used as a replacement for biopsy, that there are no incidental

**TABLE 35** The 36 scenarios analysed

| Scenario number | Biopsy sensitivity scenario | NILT accuracy scenario | Percutaneous or transjugular biopsy | Biopsy disutility |
|---|---|---|---|---|
| 1 | 100% | Scenario 1 | Percutaneous | Normal |
| 2 | 100% | Scenario 1 | Percutaneous | High |
| 3 | 100% | Scenario 1 | Transjugular | Normal |
| 4 | 100% | Scenario 1 | Transjugular | High |
| 5 | 100% | Scenario 2 | Percutaneous | Normal |
| 6 | 100% | Scenario 2 | Percutaneous | High |
| 7 | 100% | Scenario 2 | Transjugular | Normal |
| 8 | 100% | Scenario 2 | Transjugular | High |
| 9 | 100% | Scenario 3 | Percutaneous | Normal |
| 10 | 100% | Scenario 3 | Percutaneous | High |
| 11 | 100% | Scenario 3 | Transjugular | Normal |
| 12 | 100% | Scenario 3 | Transjugular | High |
| 13 | 80% pessimistic | Scenario 1 | Percutaneous | Normal |
| 14 | 80% pessimistic | Scenario 1 | Percutaneous | High |
| 15 | 80% pessimistic | Scenario 1 | Transjugular | Normal |
| 16 | 80% pessimistic | Scenario 1 | Transjugular | High |
| 17 | 80% pessimistic | Scenario 2 | Percutaneous | Normal |
| 18 | 80% pessimistic | Scenario 2 | Percutaneous | High |
| 19 | 80% pessimistic | Scenario 2 | Transjugular | Normal |
| 20 | 80% pessimistic | Scenario 2 | Transjugular | High |
| 21 | 80% pessimistic | Scenario 3 | Percutaneous | Normal |
| 22 | 80% pessimistic | Scenario 3 | Percutaneous | High |
| 23 | 80% pessimistic | Scenario 3 | Transjugular | Normal |
| 24 | 80% pessimistic | Scenario 3 | Transjugular | High |
| 25 | 80% optimistic | Scenario 1 | Percutaneous | Normal |
| 26 | 80% optimistic | Scenario 1 | Percutaneous | High |
| 27 | 80% optimistic | Scenario 1 | Transjugular | Normal |
| 28 | 80% optimistic | Scenario 1 | Transjugular | High |
| 29 | 80% optimistic | Scenario 2 | Percutaneous | Normal |
| 30 | 80% optimistic | Scenario 2 | Percutaneous | High |
| 31 | 80% optimistic | Scenario 2 | Transjugular | Normal |
| 32 | 80% optimistic | Scenario 2 | Transjugular | High |
| 33 | 80% optimistic | Scenario 3 | Percutaneous | Normal |
| 34 | 80% optimistic | Scenario 3 | Percutaneous | High |
| 35 | 80% optimistic | Scenario 3 | Transjugular | Normal |
| 36 | 80% optimistic | Scenario 3 | Transjugular | High |

benefits of conducting a biopsy, and that the results could change markedly when values are assigned to these parameters.

## Threshold analysis regarding the rates of abstinence

A threshold analysis was performed, which evaluated the change in the proportion of patients who remain abstinent required to change the conclusion so that biopsying all patients is the most cost-effective option. In this analysis, the increases were stepped in units of 0.1 percentage points until the test was no longer cost-effective, with this value reported as the threshold. For

all strategies, only the abstinence rate associated with diagnosis by the NILT was altered, with the abstinence rates associated with biopsy remaining unaffected. Thus, for a triage strategy, those patients in whom the NILT was positive would progress to biopsy and would have an abstinence rate determined by biopsy; those patients in whom the NILT was negative would have an abstinence rate associated with a negative NILT diagnosis, rather than a biopsy diagnosis, and it is the former value that is altered in the threshold analyses. For replacement strategies, both the positive and negative abstinence rates would be altered as both would be determined by the NILT alone.

This is presented for all nine NILT strategies (four triaging, four replacement and one to diagnose all with cirrhosis). The figures have been commented on to provide the reader with an aid to interpret the results.

The threshold results are presented without comment on whether or not the thresholds are realistically achievable; no data were found on abstinence rate by method of diagnosis, and any reduction based on a non-invasive diagnostic technique is a matter of genuine clinical debate. Where the threshold is a zero decrease in abstinence rates, this indicates that, for that specific scenario, biopsying all patients was deemed a more cost-effective option than the relevant strategy with equal abstinence rates and no QALY gain from biopsy. These thresholds are shown graphically in *Figures 10–27*, assuming a cost per QALY threshold of £20,000, with more detailed results presented in *Appendix 10*. The thresholds can vary markedly; it is noted that in the scenarios most favourable to the NILT, the comparator is transjugular biopsy and a high degree of disutility prior to the biopsy is assumed.

It is stressed that the threshold on abstinence is undertaken separately from the threshold on the potential benefits associated with a biopsy. Clearly the threshold for abstinence rates would decrease if there was a belief that there was also a QALY benefit associated with a biopsy.

No formal elicitation was undertaken to determine what the likely range of values for the decrease in abstinence rates when using a NILT or on the incidental benefits of a biopsy. The clinical input that was received suggested that these values would be highly uncertain.

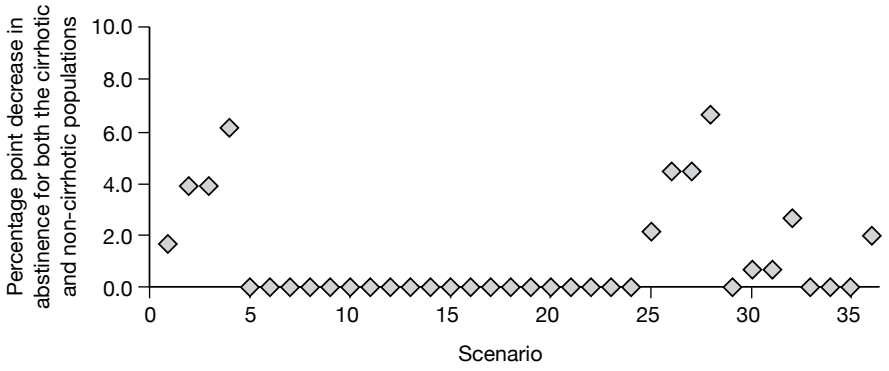



**FIGURE 10** Threshold analyses on abstinence for FibroScan triage. The threshold for change in abstinence where biopsying all becomes more cost-effective than FibroScan triage.

The scenarios that are more favourable to FibroScan triage assume scenario 1 for test accuracy. In the most favourable scenario to FibroScan triage, the abstinence rate would need to decrease by ≥ 6.7 percentage points for biopsying all to be more cost-effective; alternative scenarios estimate that biopsying all patients is more cost-effective than FibroScan triage, even assuming the same rate in abstinence.

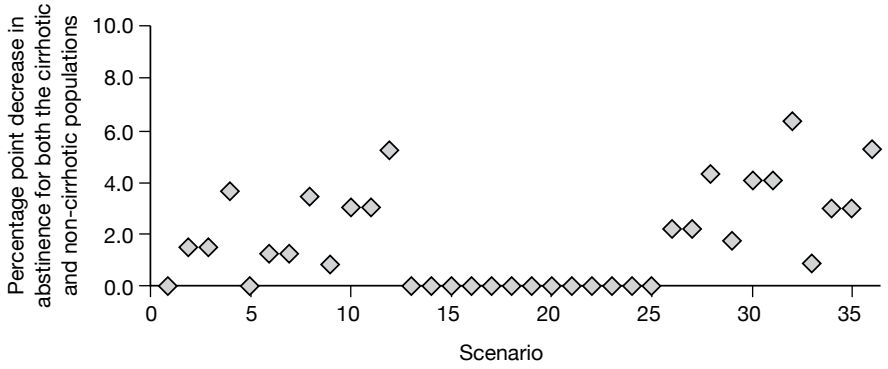



**FIGURE 11** Threshold analyses on abstinence for ELF triage. The threshold for change in abstinence where biopsying all becomes more cost-effective than ELF triage.

The ELF triage has a wide range of thresholds for decrease in abstinence levels, which can be as high as 6.3 percentage points in the most favourable scenarios, although often this value is < 2%. It is commented that the sensitivity and specificity for the ELF have not been taken from a cirrhotic population and are thus subject to more uncertainty than in the remaining tests.

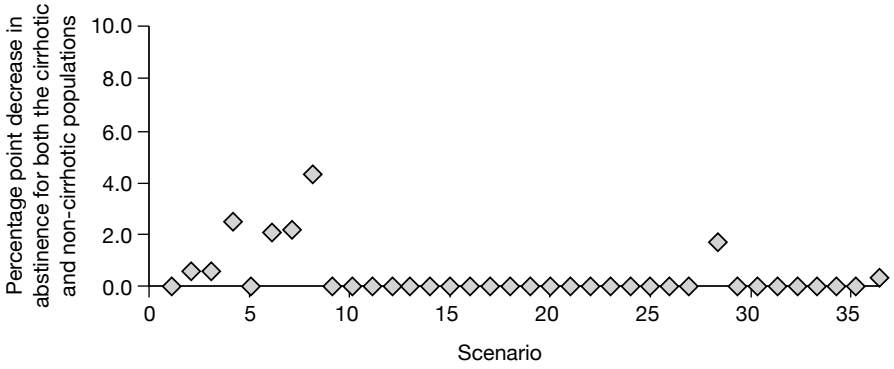



**FIGURE 12** Threshold analyses on abstinence for FibroTest triage. The threshold for change in abstinence where biopsying all becomes more cost-effective than FibroTest triage.

For all scenarios, an increase in abstinence rates by ≥ 4.3 percentage points resulted in biopsying all patients being more cost-effective than FibroTest. However, in the majority of scenarios, biopsying all patients is indicated to be more cost-effective than FibroTest triage.

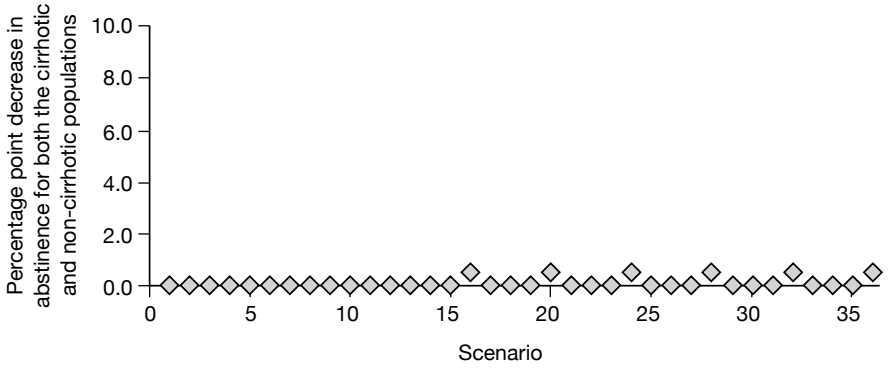



**FIGURE 13** Threshold analyses on abstinence for clinical experience triage. The threshold for change in abstinence where biopsying all becomes more cost-effective than clinical experience triage.

For all scenarios, an increase in abstinence rates of ≥ 0.5 percentage points resulted in biopsying all patients being more cost-effective than clinical experience triage. However, in the majority of scenarios, biopsying all patients is indicated to be more cost-effective than clinical experience triage.

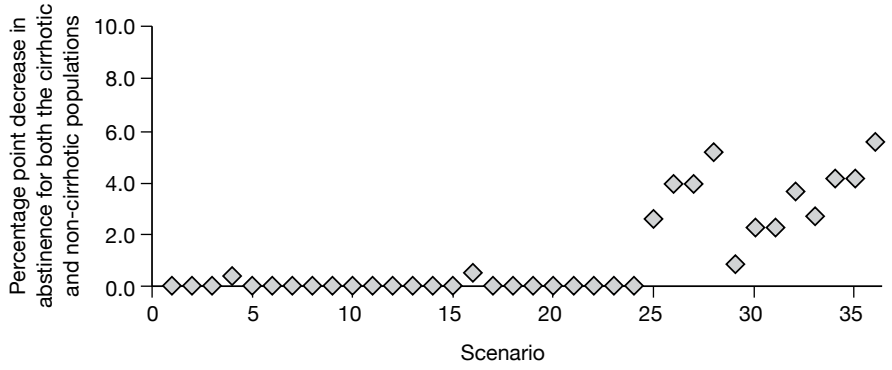



**FIGURE 14** Threshold analyses on abstinence for FibroScan replacement. The threshold for change in abstinence where biopsying all becomes more cost-effective than treating based on the FibroScan alone.

For FibroScan replacement, the threshold values only rise > 1% when it is assumed that the sensitivity of biopsy in detecting cirrhosis is 80% and that a FibroScan would have detected these FNs as being cirrhotic. In this instance, the threshold level approaches 6 percentage points.

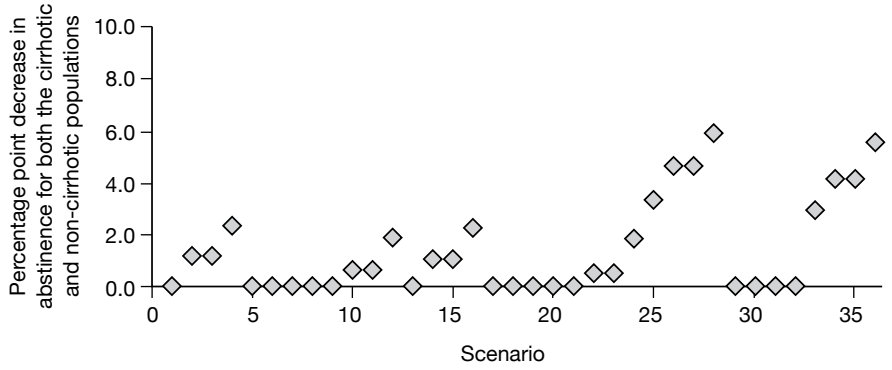



**FIGURE 15** Threshold analyses on abstinence for ELF replacement. The threshold for change in abstinence where biopsying all becomes more cost-effective than treating based on the ELF alone.

The threshold associated with ELF replacement varies markedly, but is relatively high and can reach 6 percentage points when it is assumed that the sensitivity of biopsy in detecting cirrhosis is 80% and that an ELF would have detected these FNs as being cirrhotic. It is commented that the sensitivity and the specificity for the ELF have not been taken from a cirrhotic population and are thus subject to more uncertainty than the remaining tests.

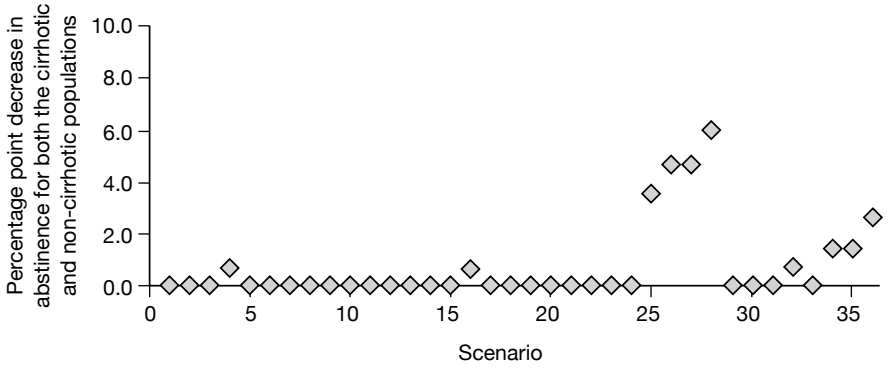



**FIGURE 16** Threshold analyses on abstinence for FibroTest replacement. The threshold for change in abstinence where biopsying all becomes more cost-effective than treating based on the FibroTest alone.

For the majority of scenarios, biopsying all patients is more cost-effective than a FibroTest replacement policy. The threshold is relatively high (reaching 6 percentage points) when it is assumed that the sensitivity of biopsy in detecting cirrhosis is 80% and that a FibroTest would have detected these FNs as being cirrhotic, and scenario 1 for FibroTest accuracy is used.

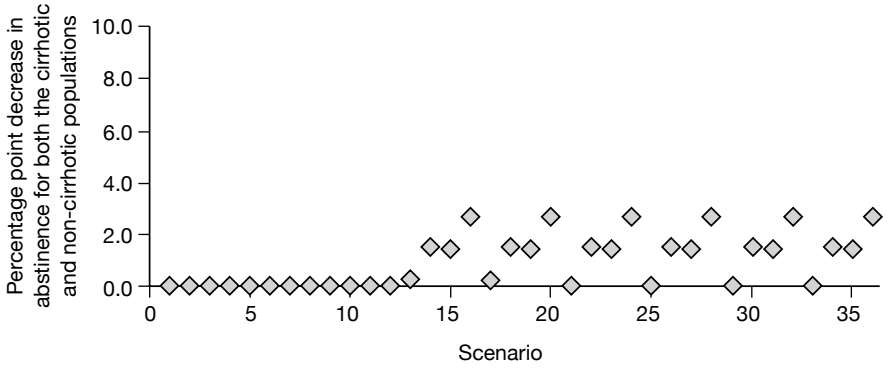



**FIGURE 17** Threshold analyses on abstinence for clinical experience replacement. The threshold for change in abstinence where biopsying all becomes more cost-effective than treating based on clinical experience alone.

For all scenarios, an increase in abstinence rates by ≥ 2.8 percentage points resulted in biopsying all patients being more cost-effective than clinical experience replacement. For a sizeable proportion of scenarios, biopsying all patients was more cost-effective than a clinical experience replacement strategy.

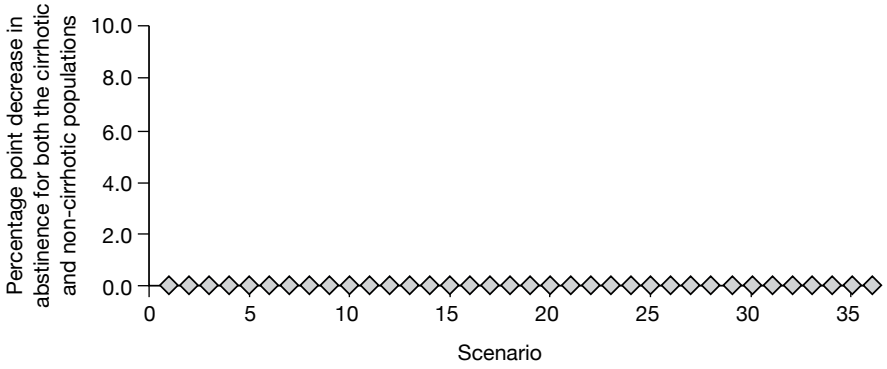



**FIGURE 18** Threshold analyses on abstinence for diagnosing all with cirrhosis replacement. The threshold for change in abstinence where biopsying all becomes more cost-effective than diagnosing all patients with cirrhosis.

For all scenarios, biopsying all patients was more cost-effective than a test that had 100% sensitivity and 0% specificity. Although the diagnose all strategy produced most QALYs, the large number of positives was associated with a much greater cost because of clinical follow-up following a diagnosis of cirrhosis.

## Threshold analysis regarding the potential benefits associated with a biopsy

A threshold analysis was performed, which evaluated the average gain in QALY from a biopsy compared with a NILT to change the conclusion so that biopsying all patients is the most cost-effective option. In this analysis, the increases were stepped in units of 0.001 QALYs until the test was no longer cost-effective, with this value reported as the threshold.

This is presented for all nine NILT strategies (four triaging, four replacement and one to diagnose all with cirrhosis). The figures have been commented on to provide the reader with an aid to interpret the results.

The threshold results are presented without comment on whether or not the thresholds are realistically achievable; no quantitative data were found on the incidental benefits of biopsy compared with a NILT. Where the threshold is a zero decrease in abstinence rates, this indicates that, for that specific scenario, biopsying all patients was deemed a more cost-effective option than the relevant strategy with equal abstinence rates and no QALY gain from biopsy. These are shown graphically in the main report (see *Figures 19–27*), with the actual values presented in *Appendix 10*.

It is stressed that the threshold on abstinence is undertaken separately from the threshold on the potential benefits associated with rates of abstinence. Thus, the threshold for QALY gain from a biopsy would decrease if there was believed to be an increase in abstinence rates following biopsy. Similarly, the threshold for abstinence rates following biopsy would decrease if there was believed to be a QALY gain from biopsy.

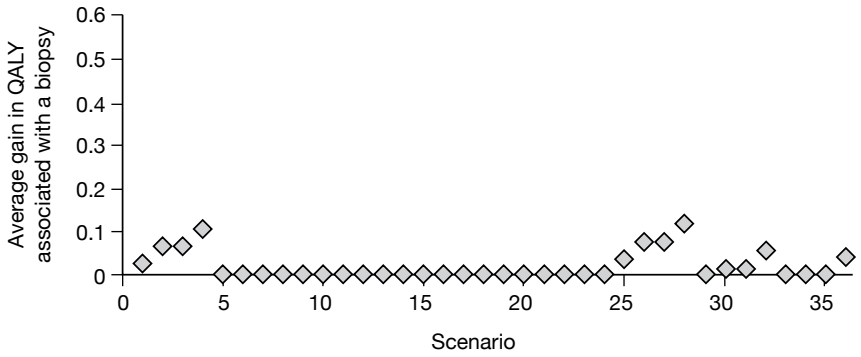



**FIGURE 19** Threshold analyses on QALY gain associated with a biopsy for FibroScan triage. The threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than FibroScan triage.

It is seen that if the QALY increase associated with a biopsy is ≥ 0.12, then biopsying all patients is the most cost-effective option. For the majority of scenarios, biopsying all patients is the most cost-effective option.

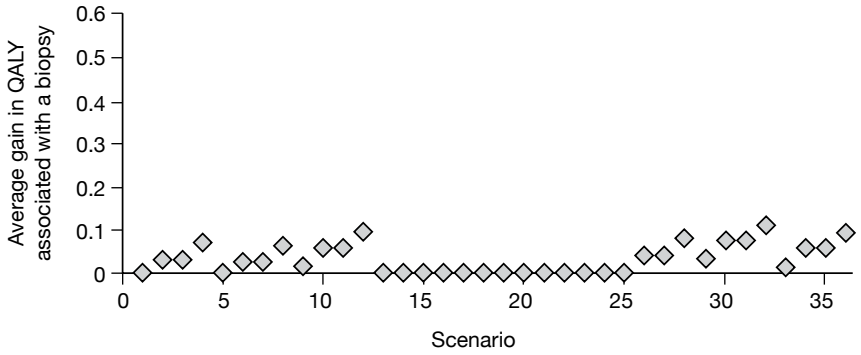



**FIGURE 20** Threshold analyses on QALY gain associated with a biopsy for ELF triage. The threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than ELF triage.

It is seen that if the QALY increase associated with a biopsy is ≥ 0.11, then biopsying all patients is the most cost-effective option. It is commented that the sensitivity and the specificity for the ELF have not been taken from a cirrhotic population and are thus subject to more uncertainty than the remaining tests.

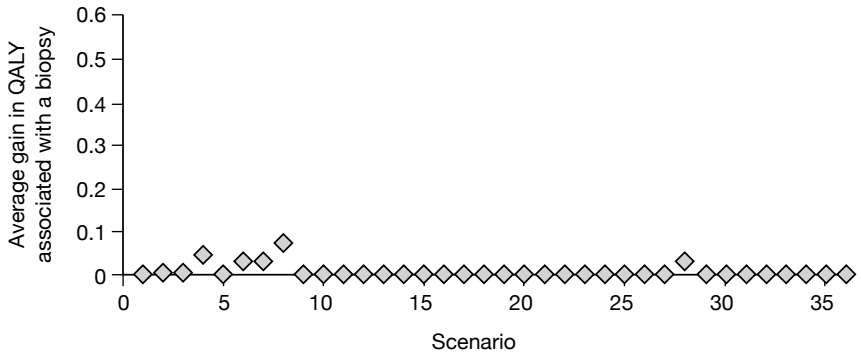



**FIGURE 21** Threshold analyses on QALY gain associated with a biopsy for FibroTest triage. The threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than FibroTest triage.

For all scenarios, a gain in biopsy of ≥ 0.08 QALYs resulted in biopsying all patients being more cost-effective than FibroTest triage. However, in the majority of scenarios, biopsying all patients is indicated to be more cost-effective than FibroTest triage.

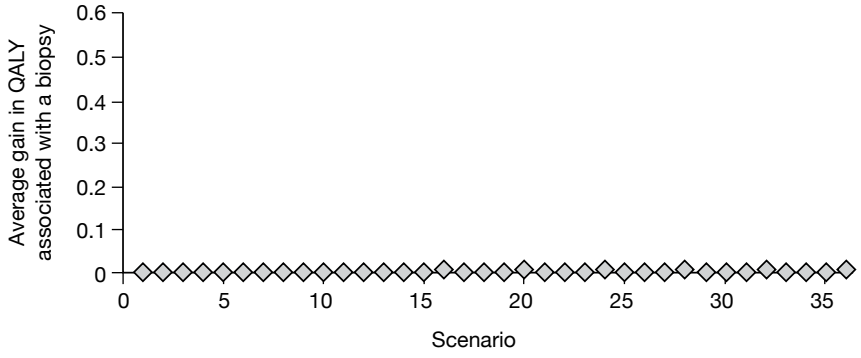



**FIGURE 22** Threshold analyses on QALY gain associated with a biopsy for clinical experience triage. The threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than clinical experience triage.

For all scenarios, a gain in biopsy of ≥ 0.01 QALYs resulted in biopsying all patients being more cost-effective than clinical experience triage. In the majority of scenarios, biopsying all patients was estimated to be more cost-effective than clinical experience triage.

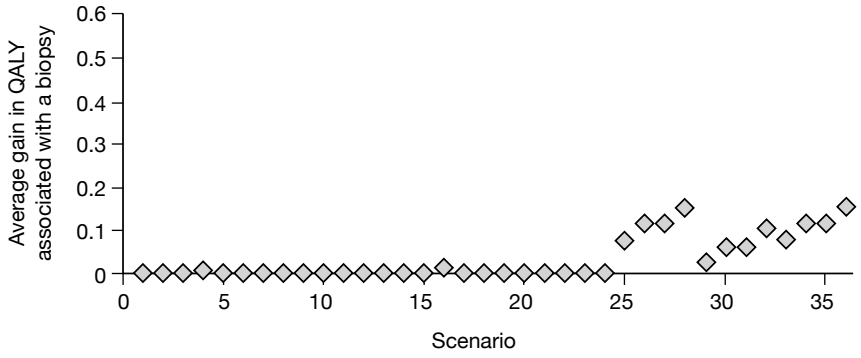



**FIGURE 23** Threshold analyses on QALY gain associated with a biopsy for FibroScan replacement. Threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than treating based on the FibroScan alone.

For FibroScan replacement, the threshold values are relatively greatest when it is assumed that the sensitivity of biopsy in detecting cirrhosis is 80% and that a FibroScan would have detected these FNs as being cirrhotic. The greatest threshold value seen is < 0.16 QALYs.

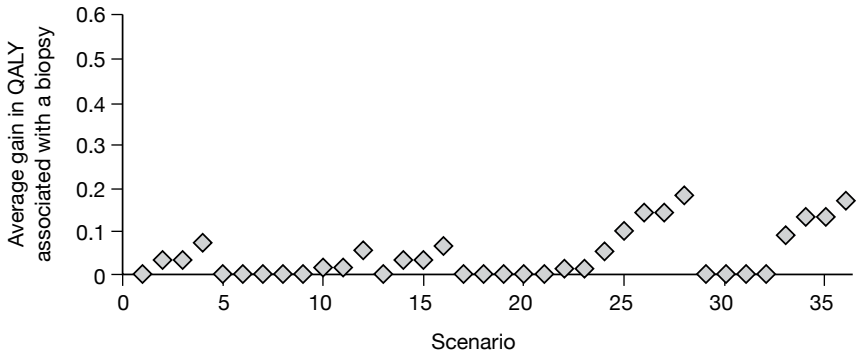



**FIGURE 24** Threshold analyses on QALY gain associated with a biopsy for ELF replacement. Threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than treating based on the ELF alone.

The greatest threshold values are produced when it is assumed that the sensitivity of biopsy in detecting cirrhosis is 80% and that an ELF would have detected these FNs as being cirrhotic. These thresholds are not estimated to be > 0.18 QALYs. It is commented that the sensitivity and the specificity for the ELF have not been taken from a cirrhotic population and are thus subject to more uncertainty than the remaining tests.

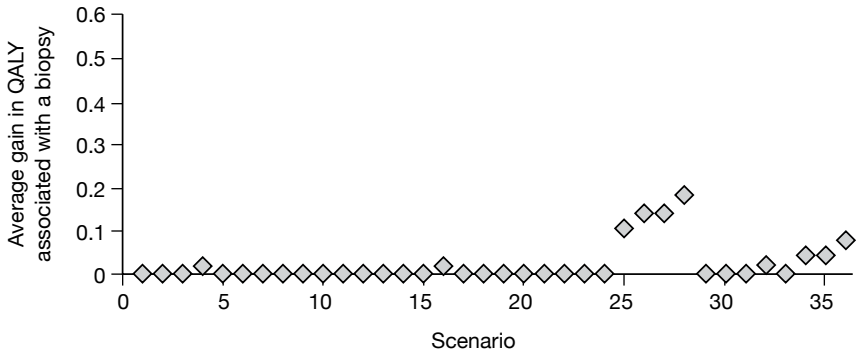



**FIGURE 25** Threshold analyses on QALY gain associated with a biopsy for FibroTest replacement. Threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than treating based on the FibroTest alone.

The results are seen to be more favourable to FibroTest replacement when it is assumed that the sensitivity of biopsy in detecting cirrhosis is 80% and that the FibroTest would have detected these FNs as being cirrhotic; in this instance, the threshold is 0.19 QALYs. However, in most scenarios, biopsying all patients is estimated to be more cost-effective than a FibroTest replacement strategy.

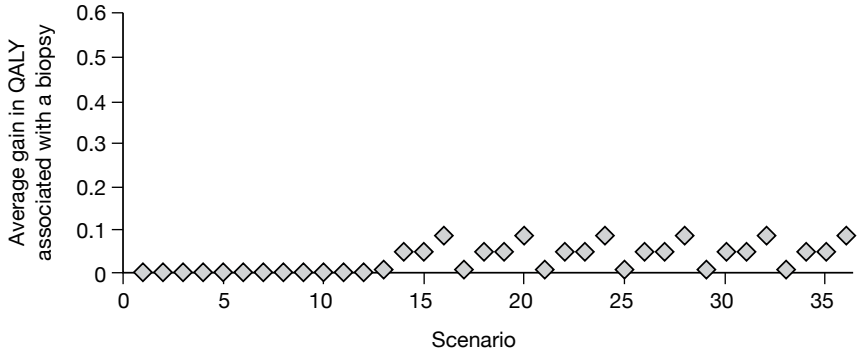



**FIGURE 26** Threshold analyses on QALY gain associated with a biopsy for clinical experience replacement. Threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than treating based on clinical experience alone.

For all scenarios, a gain in QALY of ≥ 0.09 associated with a biopsy resulted in biopsying all patients being more cost-effective than clinical experience replacement.

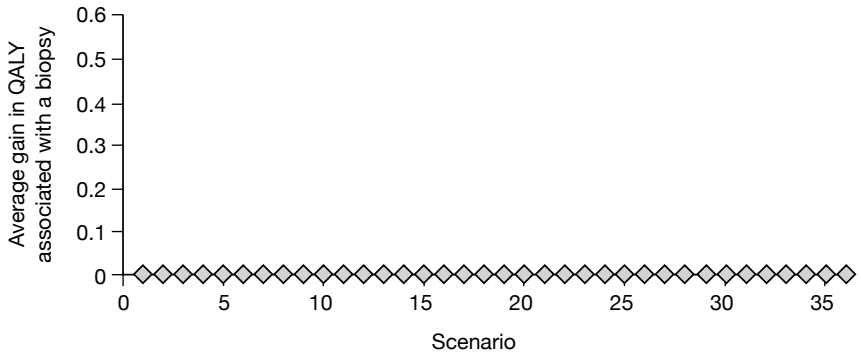



**FIGURE 27** Threshold analyses on QALY gain associated with a biopsy for diagnosing all with cirrhosis replacement. Threshold for QALY gain per biopsy performed when biopsy becomes more cost-effective than diagnosing all patients with cirrhosis.

In all scenarios, biopsying all patients was more cost-effective than a test with 100% sensitivity and 0% specificity, that is, diagnosing all with cirrhosis. Although the diagnose all strategy produced most QALYs, the large number of positives was associated with a much greater cost because of clinical follow-up following a diagnosis of cirrhosis.

## Conclusions drawn from the cost-effectiveness results

It has been shown that the cost-effectiveness of each strategy is sensitive to relatively small values in the abstinence rates assumed to be associated with that strategy. For the triage policies, in the most favourable scenarios a decrease in abstinence of ≥ 6.7 percentage points in those diagnosed as without cirrhosis results in biopsying all to be the most cost-effective strategy; this value decreases to 4.5 percentage points if scenarios in which a biopsy is associated with anxiety and thus a high disutility are excluded, and falls to 2.2 percentage points when transjugular biopsies are further excluded. In general, the threshold values were typically < 2 percentage points, and in a sizeable number of scenarios biopsying all patients was more cost-effective than a NILT triaging strategy. FibroScan and the ELF appeared to perform better than the NILTs, although the small

data set taken from a non-cirrhotic population casts more uncertainty regarding the ELF results. In some scenarios, however, FibroTest had relatively high thresholds.

For the replacement strategies, the greatest threshold for decreased abstinence rates was 6.0 percentage points, decreasing to 4.7 percentage points when scenarios in which a biopsy is associated with anxiety and thus a high disutility are excluded, and falling to 3.5 percentage points when transjugular biopsies are further excluded. FibroScan and the ELF appeared to perform better than the NILTs, although in some scenarios FibroTest had relatively high thresholds.

The conclusions from the threshold on abstinence rates also apply to the threshold analyses conducted on the QALY gains associated with a biopsy, with the cost-effectiveness results sensitive to relatively small changes in the QALY gain associated with a biopsy. For the triage policies, a QALY gain of $\geq 0.118$ in those diagnosed as without cirrhosis results in biopsying all being the most cost-effective strategy; this value decreases to 0.078 if scenarios in which a biopsy is associated with anxiety and a high disutility are excluded, and reduces to 0.038 if transjugular biopsies are excluded.

For the replacement strategies, the greatest threshold for decreased abstinence rates is 0.186 QALYs, which decreases to 0.148 QALYs when scenarios in which a biopsy is associated with anxiety and thus a high disutility are excluded, and falls to 0.108 QALYs when transjugular biopsies are further excluded.

It is stressed that the threshold analyses on decreases in abstinence rates and in potential QALY gains from biopsies have been undertaken independently, and these thresholds would fall if both a decrease in abstinence rates and a gain from a biopsy were assumed.

The sensitivity of the model to these parameters, together with the absence of data on these issues, means that no reliable estimate can be provided for either the incremental cost or the incremental QALY of a strategy when compared with biopsying all patients, and thus also the incremental cost-effectiveness ratio. Until further data are obtained, no conclusion can be provided on the most cost-effective strategy; however, there is evidence that some strategies, such as using clinical experience or diagnosing all patients with cirrhosis, will not be the most cost-effective. Accordingly, there is not sufficient evidence to support a change from current best practice.[108]

# Chapter 7

# Discussion

The estimation of the clinical effectiveness and the cost-effectiveness of NILTs for patients with suspected ALD has been difficult to conduct with precision owing to the paucity of data. As discussed in *Chapter 4, Discussion of clinical effectiveness*, there is insufficient robust evidence for the diagnostic and prognostic accuracy of NILTs in patients with suspected ALD. Moreover, even were such evidence available, there is no evidence linking test results to subsequent drinking behaviour, although long-term abstention from alcohol is known to be by far the most important management aim in patients with ALD.

The uncertainty in the clinical parameters resulted in 36 scenarios being evaluated with individual threshold analyses performed for two key variables within the conceptual model, which were the possibility of a decrease in abstinence rates associated with the NILTs compared with biopsy and the QALY gain that may be provided by a biopsy.

It is uncertain which, if any, of the 36 scenarios provide the best representation of reality, adding considerable uncertainty. The lack of data on the potential decreases in abstinence rates or QALY gains provided by a biopsy adds considerably more uncertainty, and it was seen that small changes in these values could alter the conclusion of whether or not a strategy was cost-effective compared with biopsying all patients. As such, it is not possible to provide a robust value for the incremental cost of a new strategy, the incremental QALYs of a new strategy and ultimately the incremental cost per QALY ratio. Scenarios exist in which each of the strategies analysed is more cost-effective than biopsying all patients and, in contrast, scenarios exist in which each strategy is less cost-effective than biopsying all patients.

It is plausible that patient behaviour may be affected if the specificity of the test is known, as could be the conviction with which a clinician would be able to tell the patient that they have cirrhosis if the positive predictive value of a NILT replacement strategy was low. A gain in QALYs associated with a biopsy, could also change the conclusion regarding the more cost-effective strategy when comparing biopsying all patients with diagnosing all patients with cirrhosis; if the gain per patient was > 0.19 QALYs, biopsying all patients would become the more cost-effective strategy in all scenarios.

Until better data are available for a large number of model parameters, no conclusion can be provided regarding the cost-effectiveness of NILTs in ALD. In particular, the following parameters need to be the subject of further research; however, this list does not indicate a priority order, which cannot be determined given the present limited data:

1. the sensitivity and specificity of biopsy compared with a gold standard of post-mortem assessment of fibrosis
2. the sensitivity and specificity of each NILT against a gold standard of post-mortem assessment of fibrosis (or failing that biopsy) at validated and pre-selected cut-off thresholds for the various degrees of liver damage
3. the influence of potential confounding variables, such as current drinking behaviour and the degree of hepatic inflammation, on the performance of NILTs
4. differential information on the percentage of alcohol misusers who will develop alcohol-related cirrhosis over time, by age at onset, gender and ethnic origin

5. the likelihood, and magnitude, of decreases in abstinence rates associated with a diagnosis of significant ALD by diagnostic modality compared with biopsy
6. the incidental gains in QALYs that may be associated with biopsy, owing to the determination of non-ALD-related aetiologies.

It is also noted that this report has addressed neither the issue of whether or not the provision of suitable care facilities, both before and after diagnosis, is sufficient nor the potential implications of regional variation in practice. These are key issues that would benefit from future research.

It is noted, as a limitation of the report, that study selection and data analysis were undertaken by one reviewer.

# Acknowledgements

The authors acknowledge the clinical advice provided by Dr Erika Denton, Dr Dermot Gleeson, Dr Carsten Grimm, Dr Sally Hope, Helen Manley, Gerri Mortimore, Dr Phil Newsome, Dr Stephen Pereira, Dr Helen Reeves, Professor William Rosenberg and Dr Sandy Smith. The authors also wish to thank Andrea Shippam who organised the retrieval of papers and helped in preparing and formatting the report.

## Contributions of authors

Matt Stevenson constructed the mathematical model and interpreted the results. Myfanwy Lloyd Jones performed the systematic review and interpretation of evidence. Marsha Morgan provided clinical advice throughout the project and contributed to the writing of the clinically related sections of the monograph. Ruth Wong performed the literature search. Both Myfanwy Lloyd Jones and Matt Stevenson were involved in writing the report.

# References


1. Mann RE, Smart RG, Govoni R. The epidemiology of alcoholic liver disease. *Alcohol Res Health* 2003;**27**:209–19.
2. Levitsky J, Mailliard ME. Diagnosis and therapy of alcoholic liver disease. *Semin Liver Dis* 2004;**24**:233–47.
3. Bataller R, Brenner DA. Liver fibrosis. *J Clin Invest* 2005;**115**:209–18.
4. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. *J Hepatol* 2006;**44**:217–31.
5. Office for National Statistics. *Mortality statistics. Deaths registered in 2008*. 2009. URL: http://www.ons.gov.uk/ons/rel/vsob1/mortality-statistics--deaths-registered-in-england-and-wales--series-dr-/2008/index.html (accessed 5 October 2011).
6. Patient UK. *Recommended safe limits of alcohol*. 2010. URL: http://www.patient.co.uk/health/Recommended-Safe-Limits-of-Alcohol.htm (accessed 5 October 2011).
7. Department of Health. *Sensible drinking. The report of an inter-departmental working group*. 1995. http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4084702.pdf (accessed 5 October 2011).
8. Walsh K, Alexander G. Alcoholic liver disease. *Postgrad Med J* 2000;**76**:280–6.
9. The National Clinical Guideline Centre for Acute and Chronic Conditions. *Alcohol use disorders: diagnosis and clinical management of alcohol-related physical complications. Clinical Guideline 100*. London: National Clinical Guideline Centre, The Royal College of Physicians; 2010.
10. Marsano LS, Mendez C, Hill D, Barve S, McClain CJ. Diagnosis and treatment of alcoholic liver disease and its complications. *Alcohol Res Health* 2003;**27**:247–56.
11. Fontaine H, Petitprez K, Roudot-Thoraval F, Trinchet, J-C. Guidelines for the diagnosis of uncomplicated cirrhosis. *Gastroenterol Clin Biol* 2007;**31**:504–9.
12. Jepsen P, Ott P, Andersen PK, Sørensen HT, Vilstrup H. Clinical course of alcoholic liver cirrhosis: a Danish population-based cohort study. *Hepatology* 2010;**51**:1–8.
13. Thabut D, Imbert-Bismut F, Cazals-Hatem D, Messous D, Muntenau M, Valla DC, *et al.* Relationship between the Fibrotest and portal hypertension in patients with liver disease. *Aliment Pharmacol Ther* 2007;**26**:359–68.
14. Williams R. The pervading influence of alcoholic liver disease in hepatology. *Alcohol Alcohol* 2008;**43**:393–7.
15. Goodman ZD. Grading and staging systems for inflammation and fibrosis in chronic liver diseases. *J Hepatol* 2007;**47**:598–607.
16. Standish RA, Cholongitas E, Dhillon A, Burroughs AK, Dhillon AP. An appraisal of the histopathological assessment of liver fibrosis. *Gut* 2006;**55**:569–78.
17. Friedrich-Rust M, Zeuzem S. Reproducibility and limitations of transient elastography. *Liver Int* 2009;**29**:619–20.
18. Cobbold JF, Morin S, Taylor-Robinson SD. Transient elastography for the assessment of chronic liver disease: ready for the clinic? *World J Gastroenterol* 2007;**13**:4791–7.

19. Afdhal NH, Curry M. Technology evaluation: a critical step in the clinical utilization of novel diagnostic tests for liver fibrosis. *J Hepatol* 2007;**46**:543–5.

20. Saunders JB, Walters JR, Davies AP, Paton A. A 20-year prospective study of cirrhosis. *BMJ* 1981;**282**:263–6.

21. Verrill C, Markham H, Templeton A, Carr NJ, Sheron N. Alcohol-related cirrhosis – early abstinence is a key factor in prognosis, even in the most severe cases. *Addiction* 2009;**104**:774.

22. Orrego H, Blake JE, Blendis LM, Medline A. Prognosis of alcoholic cirrhosis in the presence and absence of alcoholic hepatitis. *Gastroenterology* 1987;**92**:208–14.

23. Chedid A, Mendenhall CL, Gartside P, French SW, Chen T, Rabin L. Prognostic factors in alcoholic liver disease. *Am J Gastroenterol* 1991;**86**:210–16.

24. Cancer Research UK. *Statistics and outlook for liver cancer*. 2009. URL: http://cancerhelp.cancerresearchuk.org/type/liver-cancer/treatment/statistics-and-outlook-for-liver-cancer (accessed 5 October 2011).

25. Tome S, Lucey MR. Review article: current management of alcoholic liver disease. *Aliment Pharmacol Ther* 2004;**19**:707–14.

26. Day CP, Gilvarry E, Butler TJ, James OFW. Moderate alcohol intake is not deleterious in patients with alcoholic liver disease. *Hepatology* 1996;**24**:443A.

27. De Gottardi A, Spahr L, Gelez P, Morard I, Mentha G, Guillaud O, *et al.* A simple score for predicting alcohol relapse after liver transplantation: results from 387 patients over 15 years. *Arch Intern Med* 2007;**167**:1183–8.

28. Gleeson D. *2009 annual evidence update on alcoholic liver disease – diagnosis*. 2009. URL: http://www.library.nhs.uk/Gastroliver/ViewResource.aspx?resID=332177 (accessed 15 December 2009).

29. Department of Health; Home Office; Department for Education and Skills; Department for Culture, Media and Sport. *Safe. Sensible. Social. The next steps in the National Alcohol Strategy*. 2007. URL: http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_075219.pdf (accessed 5 October 2011).

30. Fuller E. Adult alcohol consumption. In Craig R, Shelton N, editors. *Health Survey for England. Volume 1 Healthy lifestyles: knowledge, attitudes and behaviour*. London: The Information Centre for Health and Social Care; 2008. pp. 177–218.

31. Fuller E, Jotangia D, Farrell M. Alcohol misuse and dependence. In McManus S, Meltzer H, Brugha T, Bebbington P, Jenkins R, editors. *Adult psychiatric morbidity in England, 2007. Results of a household survey*. London: The Health and Social Care Information Centre; 2009. pp. 151–173.

32. Statistics for Wales. *Welsh health survey 2008*. 2009. URL: http://wales.gov.uk/topics/statistics/publications/publication-archive/healthsurvey2008/?lang=en (accessed 5 October 2011).

33. Gharial N. Alcohol as a problem for the south Asian community. *UK Alcohol Alert* 2007;**1**:21–2.

34. Douds AC, Cox MA, Iqbal TH, Cooper BT. Ethinic differences in cirrhosis of the liver in a British city: alcoholic cirrhosis in South Asian men. *Alcohol Alcohol* 2003;**38**:148–50.

35. British Liver Trust. *Fighting liver disease: facts about liver disease*. 2009. URL: http://www.britishlivertrust.org.uk/home/about-us/media-centre/facts-about-liver-disease.aspx (accessed 10 December 2009).

36. Raynard B, Balian A, Fallik D, Capron F, Bedossa P, Chaput JC, *et al.* Risk factors of fibrosis in alcohol-induced liver disease. *Hepatology* 2002;**35**:635–8.

37. Hart CL, Morrison DS, Batty GD, Davey Smith G. Effect of body mas index and liver consumption on liver disease: analysis of data from two prospective cohort studies. *BMJ* 2010; **340**:c1240.

38. The information centre for health and social care. *Health Survey for England – 2008. Adult trend tables*. 2009. URL: http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles-related-surveys/health-survey-for-england/health-survey-for-england--2008-trend-tables (accessed 5 July 2010).

39. National Heart Forum. *Type 2 diabetes*. 2006. URL: http://www.heartforum.org.uk/AboutCHD_riskfac_type2diab.aspx (accessed 5 July 2010).

40. Fleming KM, Aithal GP, Solaymani-Dodaran M, Card TR, West J. Incidence and prevalence of cirrhosis in the United Kingdom, 1992–2001: a general population-based study. *J Hepatol* 2008;**49**:732–8.

41. Schuppan D, Afdhal NH. Liver cirrhosis. *Lancet* 2008;**371**:838–51.

42. de Franchis R, Dell'Era A, Primignani M. Diagnosis and monitoring of portal hypertension. *Dig Liver Dis* 2008;**40**:312–17.

43. Herrera JL, Rodriguez R. Medical care of the patient with compensated cirrhosis. *Gastroenterol Hepatol* 2006;**2**:124–33.

44. Sandhu BS, Sanyal AJ. Management of ascites in cirrhosis. *Clin Liver Dis* 2005;**9**:715–732.

45. Sargent S. The management and nursing care of cirrhotic ascites. *Br J Nurs* 2006;**15**:212–19.

46. Amodio P, Montagnese S, Gatta A, Morgan MY. Characteristics of minimal hepatic encephalopathy. *Metab Brain Dis* 2004;**19**:253–67.

47. Poordad FF. Review article: the burden of hepatic encephalopathy. *Aliment Pharmacol Ther* 2007;**25**:3–9.

48. NHS Information Centre. *Statistics on alcohol: England, 2009*. London: The Health and Social Care Information Centre; 2009.

49. Coles EC. *Alcohol and health in Wales: a major public health issue*. Cardiff: National Public Health Service for Wales; 2006.

50. Welch C, Harrison D, Short A, Rowan K. The increasing burden of alcoholic liver disease on United Kingdom critical care units: secondary analysis of a high quality clinical database. *J Health Serv Res Policy* 2008;**13**:40–4.

51. National Institute for Health Research National Horizon Scanning Centre. *Enhanced Liver Fibrosis Test (ELF) for evaluating liver fibrosis*. Birmingham: National Horizon Scanning Centre; 2008.

52. Stewart SF, Day CP. The management of alcoholic liver disease. *J Hepatol* 2003;**38**:S2–S13.

53. Rambaldi A, Saconato HH, Christensen E, Thorlund K, Wetterslev J, Gluud C. Systematic review: glucocorticosteroids for alcoholic hepatitis – a Cochrane Hepato-Biliary Group systematic review with meta-analyses and trial sequential analyses of randomized clinical trials. *Aliment Pharmacol Ther* 2008;**27**:1167–78.

54. Thabut D, Moreau R, Lebrec D. Screening for esophageal varices: endoscopy, other tools, or endoscopy and other tools? *Hepatology* 2008;**47**:1434–6.

55. Sotiropoulos GC, Beckebaum S, Lang H, Frilling A, Molmenti EP, Cicinnati VR, *et al.* Single-center experience on liver transplantation for hepatocellular carcinoma arising in alcoholic cirrhosis: results and ethical issues. *Eur Surg Res* 2008;**40**:7–13.

56. Stamuli E, Kruger J, Hutton J. *Cost-effectiveness of ultrasound elastography in the assessment of liver fibrosis*. Centre for Evidence-Based Purchasing Economic Report CEP08053; 2009.

57. Parkes J, Roderick P, Harris S, Day C, Mutimer D, Collier J, *et al.* Enhanced liver fibrosis test can predict clinical outcomes in patients with chronic liver disease. *Gut* 2010;**59**:1245–51.

58. iQur Ltd. *Enhanced Liver Fibrosis Test (ELFTM Test).* 2006. URL: http://www.iqur.com/Sampleprep.html (accessed 22 December 2009).

59. Campbell S, Timms PM, Maxwell PR, Doherty EM, Rahman MZ, Lean ME, *et al.* Effect of alcohol withdrawal on liver transaminase levels and markers of liver fibrosis. *J Gastroenterol Hepatol* 2001**;16**:1254–9.

60. Tsutsumi M, Urashima S, Takase S, Ueshima Y, Tsuchishima M, Shimanaka K, *et al.* Characteristics of serum hyaluronate concentrations in patients with alcoholic liver disease. *Alcohol Clin Exp Res* 7010;**21**:1716–21.

61. Nouchi T, Worner TM, Sato S, Lieber CS. Serum procollagen type III N-terminal peptides and laminin Pl peptide in slcoholic liver disease. *Alcohol Clin Exp Res* 1987;**11**:7–287.

62. Lieber CS, Weiss DG, Paronetto F, Veterans Affairs Cooperative Study 391 Group. Value of fibrosis markers for staging liver fibrosis in patients with precirrhotic alcoholic liver disease. *Alcohol Clin Exp Res* 2008;**32**:1031–9.

63. Castera L, Pinzani M. Biopsy and non-invasive methods for the diagnosis of liver fibrosis: does it take two to tango? *Gut* 2010;**59**:861–6.

64. National Institute for Health Research National Horizon Scanning Centre. *Fibro-Test-ActiTest™ for diagnosis and monitoring of fibrosis in chronic liver conditions.* Birmingham: National Horizon Scannning Centre; 2005.

65. Poynard T, Halfon P, Castera L, Charlotte F, Le Bail B, Munteanus M, *et al.* Variability of the area under the receiver operating characteristic curves in the diagnostic evaluation of liver fibrosis markers: impact of biopsy length and fragmentation. *Aliment Pharmacol Ther* 2007;**25**:733–9.

66. Rosenthal-Allieri MA, Peritore ML, Tran A, Halfon P, Benzaken S, Bernard A. Analytical variability of the fibrotest proteins. *Clin Biochem* 2005;**38**:473–8.

67. Morra R, Munteanu M, Imbert-Bismut F, Messous D, Ratziu W, Poynard T. FibroMAXTM: towards a new universal biomarker of liver disease? *Expert Rev Mol Diagn* 2007;**7**:481–90.

68. BioPredictive. *Practice of FibroMax for alcoholic liver disease*. 2010. URL: http://www.biopredictive.com/intl/physician/physicians/fibromax-for-alcohol/ (accessed 8 July 2010).

69. BioPredictive. *Let's assess our liver*. 2009. URL: http://www.biopredictive.com/intl/patient (accessed 22 December 2009).

70. Halfon P, Imbert-Bismut F, Messous D, Antoniotti G, Benchetrit D, Cart-Lamy P, *et al.* A prospective assessment of the inter-laboratory variability of biochemical markers of fibrosis (FibroTest) and activity (ActiTest) in patients with chronic liver disease. *Comp Hepatol* 2002;**1**:3.

71. Calès P, Veillon P, Konaté A, Mathieu E, Ternisien C, Chevailler A, *et al.* Reproducibility of blood tests of liver fibrosis in clinical practice. *Clin Biochem* 2008;**41**:10–18.

72. Gressner OA, Beer N, Jodlowski A, Gressner AM. Impact of quality control accepted inter-laboratory variations on calculated Fibrotest/Actitest scores for the non-invasive biochemical assessment of liver fibrosis. *Clin Chim Acta* 2009;**409**:90–5.

73. Poynard T, Munteanu M, Imbert-Bismut F, Charlotte F, Thabut D, Le Calvez S, *et al.* Prospective analysis of discordant results between biochemical markers and biopsy in patients with chronic hepatitis C. *Clin Chem* 2004;**50**:1344–55.

74. MedlinePlus. *Bilirubin – blood*. 2009., URL: http://www.nlm.nih.gov/medlineplus/ency/article/003479.htm (accessed 5 January 2010).

75. National Institute for Health Research National Horizon Scanning Centre. *Transient elastography (FibroScan) for evaluating liver fibrosis*. Birmingham: National Horizon Scanning Centre; 2008.

76. Echosens. *FibroScan*. 2009. URL: http://www.echosens.com/pdf/3en.pdf (accessed 22 December 2009).

77. Ziol M, Handra-Luca A, Kettaneh A, Christidis C, Mal F, Kazemi F, *et al.* Noninvasive assessment of liver fibrosis by measurement of stiffness in patients with chronic hepatitis C. *Hepatology* 2005;**41**:48–54.

78. Poynard T, Ingiliz P, Elkrief L, Munteanu M, Lebray P, Morra R, *et al.* Concordance in a world without a gold standard: a new non-invasive methodology for improving accuracy of fibrosis markers. *PLoS ONE* 2008;**3**:e3857.

79. Castéra L, Foucher J, Bernard PH, Carvalho F, Allaix D, Merrouche W, *et al.* Pitfalls of liver stiffness measurement: a 5-year prospective study of 13,369 examinations. *Hepatology* 2010;**51**:828–35.

80. Roulot D, Czernichow S, Le Clésiau H, Costes JL, Vergnaud AC, Beaugrand M. Liver stiffness values in apparently healthy subjects: influence of gender and metabolic syndrome. *J Hepatol* 2008;**48**:606–13.

81. Castera L, Vergniol J, Foucher J, Le Bail B, Chanteloup E, Haaser M, *et al.* Prospective comparison of transient elastography, Fibrotest, APRI, and liver biopsy for the assessment of fibrosis in chronic hepatitis C. *Gastroenterology* 2005;**128**:343–50.

82. Castera L, Forns X, Alberti A. Non-invasive evaluation of liver fibrosis using transient elastography. *J Hepatol* 2008;**48**:835–47.

83. Mederacke I, Wursthorn K, Kirschner J, Rifai K, Manns MP, Wedemeyer H, *et al.* Food intake increases liver stiffness in patients with chronic or resolved hepatitis C virus infection. *Liver Int* 2009;**29**:1500–6.

84. Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, Ronchi G, *et al.* Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. *Gut* 2007;**56**:968–73.

85. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, Mal F, *et al.* Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Bio* 2003;**29**:1705–13.

86. Bravo AA, Sheth SG, Chopra S. Liver biopsy. *N Engl J Med* 2001;**344**:495–500.

87. Patient UK. *Liver biopsy*. 2010. URL: http://www.patient.co.uk/health/Biopsy-Liver.htm (accessed 29 April 2010).

88. Neuberger J, Grant A, Day C, Saxseena S. *Guidelines on the use of liver biopsy in clinical practice*. British Society of Gastroenterology. 2004;1–15. URL: http://www.bsg.org.uk/images/stories/docs/clinical/guidelines/liver/liver_biopsy.pdf (accessed 29 April 2010).

89. Regev A, Berho M, Jeffers LJ, Milikowski C, Molina EG, Pyrsopoulos NT, *et al.* Sampling error and intraobserver variation in liver biopsy in patients with chronic HCV infection. *Am J Gastroenterol* 2002;**97**:2614–18.

90. Bedossa P, Dargère D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003;**38**:1449–57.

91. Abdi W, Millan JC, Mezey E. Sampling variability on percutaneous liver biopsy. *Arch Intern Med* 1979;**139**:667–9.

92. Maharaj B, Maharaj RJ, Leary WP. Sampling variability and its influence on the diagnostic yield of percutaneous needle biopsy of the liver. *Lancet* 1986;**1**:523–5.

93. Murtagh J, Foerster V. Transient elastography (FibroScan) for non-invasive assessment of liver fibrosis. *Issues Emerg Health Technol* 2006;**90**:1–4.

94. The French METAVIR Cooperative Study Group. Intraobserver and interobserver variations liver biopsy interpretation in patients with chronic hepatitis C. *Hepatology* 1994;**20**:15–20.

95. Plebani M, Basso D. Non-invasive assessment of chronic liver and gastric diseases. *Clin Chim Acta* 2007;**381**:39–49.

96. Van Thiel DH, Gavaler JS, Wright H, Tzakis A. Liver biopsy. Its safety and complications as seen at a liver transplant center. *Transplantation* 1993;**55**:1087–90.

97. Lemoine M, Katsahian S, Ziol M, Nahon P, Ganne-Carrie N, Kazemi F, *et al.* Liver stiffness measurement as a predictive tool of clinically significant portal hypertension in patients with compensated hepatitis C virus or alcohol-related cirrhosis. *Aliment Pharmacol Ther* 2008;**28**:1102–10.

98. Kumar A, Sharma P, Sarin SK. Hepatic venous pressure gradient measurement: time to learn! *Indian J Gastroenterol* 2008;**27**:74–80.

99. Thabut D, Trabut JB, Massard J, Rudler M, Muntenau M, Messous D, *et al.* Non-invasive diagnosis of large oesophageal varices with FibroTest in patients with cirrhosis: a preliminary retrospective study. *Liver Int* 2006;**26**:271–8.

100. Bosch J. Predictions from a hard liver. *J Hepatol* 2006;**45**:174–7.

101. Eisen G, Baron TH, Dominitz J, Faigel DO, Goldstein JL, Johanson JF, *et al.* Complications of upper GI endoscopy. *Gastrointest Endosc* 2002;**55**:784–93.

102. Artemis Medical. *Echosens Fibroscan. Current installations in Great Britain*. 2009. URL: http://www.artemismedical.co.uk/fibroscan_installations.htm (accessed 29 April 2010).

103. Deeks JJ. Systematic reviews of evaluations of diagnostic and screening tests. In Egger M, Davey Smith G, Altman DG, editors. *Systematic reviews in health care: meta-analysis in context*. 2nd edn. London: BMJ Books; 2001. pp. 248–284.

104. Bedossa P, Carrat F. Liver biopsy: the best, not the gold standard. *J Hepatol* 2009;**50**:1–3.

105. Lambert J, Halfon P, Penaranda G, Bedossa P, Cacoub P, Carrat F. How to measure the diagnostic accuracy of noninvasive liver fibrosis indices: the area under the ROC curve revisited. *Clin Chem* 2008;**54**:1372–8.

106. Poynard T, Halfon P, Castera L, Munteanu M, Imbert-Bismut F, Ratziu V, *et al.* FibroPaca Group Standardization of ROC curve areas for diagnostic evaluation of liver fibrosis markers based on prevalences of fibrosis stages. *Clin Chem* 2007;**53**:1615–22.

107. Pinzani M, Rombouts K, Colagrande S. Fibrosis in chronic liver diseases: diagnosis and management. *J Hepatol* 2005;**42**:S22–S36.

108. National Institute for Health and Clinical Excellence (NICE). *Alcohol use disorders: diagnosis, assessment and management of harmful drinking and alcohol dependence.* Full guideline draft for consultation June 2010. 2010. URL: http://www.nice.org.uk/nicemedia/live/11875/49448/49448.pdf (accessed 7 July 2010).

109. Longworth L, Young T, Buxton MJ, Ratcliffe J, Neuberger J, Burroughs A, *et al.*, CELT Project Team Midterm cost-effectiveness of the liver transplantation program of England and Wales for three disease groups. *Liver Transpl* 2003;**9**:1295–307.

110. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet* 1999;**354**:1896–900.

111. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:332–9.

112. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.

113. Friedrich-Rust M, Ong M, Martens S, Sarrazin C, Bojunga J, Zeuzem S, *et al.* Performance of transient elastography for the staging of liver fibrosis: a meta-analysis. *Gastroenterology* 2008;**134**:960–74.

114. Berlin JA. Does blinding of readers affect the results of meta-analyses? *Lancet* 1997;**350**:185–6.

115. Clark HD, Wells JA, Huet C, McAlister FA, Salmi LR, Fergusson D, *et al.* Assessing the quality of randomized trials: reliability of the Jadad scale. *Clin Trials* 1999;**20**:448–52.

116. Deeks JJ. Systematic reviews in heatlh care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;**323**:157–62.

117. Janssens F, de Suray N, Piessevaux H, Horsmans Y, de Timary P, Starkel P. Can transient elastography replace liver histology for determinatiaon of advanced fibrosis in alcoholic patients: a real-life study. *J Clin Gastroenterol* 2010;**44**:S75–82.

118. Mueller S, Millonig G, Sarovska L, Friedrich S, Riemann FM, Pritsch M, *et al.* Increased liver stiffness in alcoholic liver disease: differentiating fibrosis from steatohepatitis. *World J Gastroenterol* 2010;**16**:966–72.

119. Janssens F, de Surray N, Horsmans Y, Piessevaux H, de Timary P, Starkel P. Determination of severe fibrosis and cirrhosis in alcoholic patients by transient elastography: a prospective comparison with liver biopsy. *J Hepatol* 2009;**50**:S362.

120. Mueller S, Millonig G, Friedrich S, Welker A, Becker P, Reimann F, *et al.* Diagnosis of cirrhosis in alcoholic liver disease (ALD): is there a place for transient elastography? *Gastroenterology* 2008;**134**:A307–8.

121. Rosenberg WM, Voelker M, Thiel R. Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology* 2004;**127**:1704–13.

122. Melin P, Schoeny M, Dacon A, Gauchet A, Diebold MD. Dépistage non invasif de la fibrose hépatique. Intérêt du FibroScan(R) en consultation d'alcoologie. *Alcoologie et Addictologie* 2005;**27**:191–6.

123. Naveau S, Raynard B, Ratziu V, Abella A, Imbert-Bismut F, Messous D, *et al.* Biomarkers for the prediction of liver fibrosis in patients with chronic alcoholic liver disease. *Clin Gastroenterol Hepatol* 2005;**3**:167–74.

124. Naveau S, Gaudé G, Asnacios A, Agostini H, Abella A, Barri-Ova N, *et al.* Diagnostic and prognostic values of noninvasive biomarkers of fibrosis in patients with alcoholic liver disease. *Hepatology* 2009;**49**:97–105.

125. Thabut D, Trabut JB, Le Calvez S, Thibaut V, Massard J, d'Arondel C, *et al.* Diagnostic value of fibrosis biochemical markers (FibroTest) for the screening of oesophageal varices in patients with chronic liver disease. *Hepatology* 2003;**38**:293A.

126. Thabut D, Lebrec D, Oberti F, Perarnau J, Condat B, Barraud H, *et al.* Prognostic value of algorithms combining FibroTest®FibroSURE® (FT) and AshTest® (HT) in comparison with MELD and Pugh prognostic indexes in patients with severe cirrhosis. *Hepatology* 2007;**46**:594A.

127. Kim SG, Kim YS, Jung SW, Kim HK, Jang JY, Moon JH, *et al.* The usefulness of transient elastography to diagnose cirrhosis in patients with alcoholic liver disease. *Korean J Hepatol* 2009;**15**:42–51.

128. Nahon P, Kettaneh A, Tengher-Barna I, Ziol M, de Ledinghen V, Douvin C, *et al.* Assessment of liver fibrosis using transient elastography in patients with alcoholic liver disease. *J Hepatol* 2008;**49**:1062–8.

129. Bureau C, Metivier S, Peron JM, Selves J, Robic MA, Gourraud PA, *et al.* Transient elastography accurately predicts presence of significant portal hypertension in patients with chronic liver disease. *Aliment Pharmacol Ther* 2008;**27**:1261–8.

130. Nguyen-Khac E, Saint-Leger P, Tramier B, Coevoet H, Capron D, Dupas JL. Non-invasive diagnosis of large esophageal varices by Fibroscan: strong influence of the cirrhosis aetiology. *Hepatology* 2009;**50**:742A.

131. Nahon P, Ziol M, de Ledinghen V, Douvin C, Marcellin P, Ganne-Carrie N, *et al.* Assessment of liver fibrosis using liver stiffness measurement (LSM) in patients with alcoholic liver disease. *J Hepatol* 2007;**46**:S278–S279.

132. Nguyen-Khac E, Chatelain D, Tramier B, Decrombecque C, Robert B, Joly JP, *et al.* Assessment of asymptomatic liver fibrosis in alcoholic patients using fibroscan: prospective comparison with seven non-invasive laboratory tests. *Aliment Pharmacol Ther* 2008;**28**:1188–98.

133. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. *Hepatology* 1996;**24**:289–93.

134. Batts KP, Ludwig J. Chronic hepatitis. An update on terminology and reporting. *Am J Surg Pathol* 2010;**19**:1409–17.

135. Chevallier M, Guerret S, Chossegros P, Gerard F, Grimaud JA. A histological semiquantitative scoring system for evaluation of hepatic fibrosis in needle liver biopsy specimens: comparison with morphometric studies. *Hepatology* 1994;**20**:349–355.

136. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005;**41**:1313–21.

137. Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol* 1999;**94**:2474.

138. Tsochatzis EA, Gurusamy KS, Ntaoula S, Cholongitas E, Davidson BR, Burroughs AK. Elastography for the diagnosis of severity of fibrosis in chronic liver disease: a meta-analysis of diagnostic accuracy. *J Hepatol* 2010;**54**:650–9.

139. Kettaneh A, Marcellin P, Douvin C, Poupon R, Ziol M, Beaugrand M, *et al.* Features associated with success rate and performance of FibroScan measurements for the diagnosis of cirrhosis in HCV patients: a prospective study of 935 patients. *J Hepatol* 2007;**46**:628–34.

140. de Ledinghen V, Fournier C, Foucher J, Miette V, Vergniol J, Rigalleau V, *et al.* New FibroScan probe for obese patients. A pilot study of feasibility and performances in patients with BMI ≥ 30 kg/m$^2$. *J Hepatol* 2009;**50**:989.

141. Sagir A, Erhardt A, Schmitt M, Häussinger D. Transient elastography is unreliable for detection of cirrhosis in patients with acute liver damage. *Hepatology* 2008;**47**:592–5.

142. Arena U, Vizzutti F, Corti G, Ambu S, Stasi C, Bresci S, *et al.* Acute viral hepatitis increases liver stiffness values measured by transient elastography. *Hepatology* 2008;**47**:380–4.

143. Ghany MG, Doo E. Assessment of liver fibrosis: palpate, poke or pulse? *Hepatology* 2005;**42**:759–61.

144. Bosch J. Towards the non-invasive diagnosis of cirrhosis: the nuts-cirrhosis connection. *J Hepatol* 2009;**50**:4–6.

145. Melin P, Dacon A, Gauchet A, Schoeny M, Fournier C, Sandrin L, *et al.* Cirrhosis screening in alcoholism consultation using FibroScan (R). *Hepatology* 2005;**42**:492A–3A.

146. Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ* 1994;**309**:102.

147. van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ. Diagnostic methods I: sensitivity, specificity, and other measures of accuracy. *Kidney Int* 2009;**75**:1257–63.

148. Mehta SH, Lau B, Afdhal NH, Thomas DL. Exceeding the limits of liver histology markers. *J Hepatol* 2009;**50**:36–41.

149. Rosenberg WM. Rating fibrosis progression in chronic liver diseases. *J Hepatol* 2003;**38**:357–60.

150. National Institute for Health and Clinical Excellence. *Guide to the methods of technology appraisals*. 2008. URL: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf (accessed October 2009).

151. Carlson JJ, Kowdley KV, Sullivan SD, Ramsey SD, Veenstra DL. An evaluation of the potential cost-effectiveness of non-invasive testing strategies in the diagnosis of significant liver fibrosis. *J Gastroenterol Hepatol* 2009;**24**:786–91.

152. Curtis L. *Unit costs of health and social care*. Canterbury:PSSRU, University of Kent; 2009.

153. Thompson Coon J, Rogers G, Hewson P, Wright D, Anderson R, Jackson S, *et al.* Surveillance of cirrhosis for hepatocellular carcinoma: a cost-utility analysis. *Br J Cancer* 2008;**98**:1166–75.

154. Wright M, Main J, Goldin RD, Thomas HC. Health benefits from anti viral therapy for mild chronic hepatitis C. A UK multi-centre randomised controlled trial. *Hepatology* 2003;**38**:330.

155. Department of Health. *NHS reference costs 2007–08*. London: Department of Health; 2009.

156. Huang E, Esrailian E, Spiegel BM. The cost-effectiveness and budget impact of competing therapies in hepatic encephalopathy – a decision analysis. *Aliment Pharmacol Ther* 2007;**26**:1147–61.

157. Kind P, Hardman G, Macran S. *UK population norms for EQ-5D*. York: Centre for Health Economics, University of York; 1999.

158. de Franchis R. Revising consensus in portal hypertension: report of the Baveno V consensus workshop on methodology of diagnosis and therapy in portal hypertension. *J Hepatol* 2010;**53**:762–8.

159. Arguedas MR, Heudebert GR, Eloubeidi MA, Abrams GA, Fallon MB. Cost-effectiveness of screening, surveillance, and primary prophylaxis strategies for esophageal varices. *Am J Gastroenterol* 2002;**97**:2441–52.

160. British Medical Association and Royal Pharmaceutical Society of Great Britain. *British national formulary*. No. 60. September 2010. London: BMA and RPS; 2010.

161. Thabut D, Hammer M, Cai Y, Carbonell N. Cost of treatment of oesophageal variceal bleeding in patients with cirrhosis in France: results of a French survey (brief record). *Eur J Gastroenterol Hepatol* 2007;**19**:679–86.

162. Wechowski J, Connolly M, Woehl A, Tetlow A, McEwan P, Burroughs A, *et al.* An economic evaluation of vasoactive agents used in the United Kingdom for acute bleeding oesophageal varices in patients with liver cirrhosis. *Curr Med Res Opin* 2007;**23**:1481–91.

163. Claxton K, Sculpher M, McCabe C, Briggs A, Akehurst R, Buxton M, *et al.* Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. *Health Economics* 2005;**14**:339–47.

164. Knodell RG, Ishak KG, Black WC, Chen TS, Craig R, Kaplowitz N, *et al.* Formulation and application of a numerical scoring system for assessing histological activity in asymptomatic chronic active hepatitis. *Hepatology* 1981;**1**:431–5.

165. Ishak K, Baptista A, Bianchi L, Callea F, De Groote J, Gudat F, *et al.* Histological grading and staging of chronic hepatitis. *J Hepatol* 1995;**22**:696–9.

166. Scheuer PJ. Classification of chronic viral hepatitis. A need for reassessment. *J Hepatol* 1991;**13**:372–4.

167. Beddy P, Lyburn IL, Geoghegan T, Buckley O, Buckley AR, Torreggiani WC. Outpatient liver biopsy: a prospective evaluation of 500 cases. *Gut* 2007;**56**:307.

168. Caturelli E, Giacobbe A, Facciorusso D, Bisceglia M, Villani MR, Siena DA, *et al.* Percutaneous biopsy in diffuse liver disease: increasing diagnostic yield and decreasing complication rate by routine ultrasound assessment of puncture site. *Am J Gastroenterol* 1996;**91**:1318–21.

169. Douds AC, Joseph AE, Finlayson C, Maxwell JD. Is day case liver biopsy underutilised? *Gut* 1995;**37**:574–5.

170. Firpi RJ, Soldevila PC, Abdelmalek MF, Morelli G, Judah J, Nelson DR. Short recovery time after percutaneous liver biopsy: should we change our current practices? *Clin Gastroenterol Hepatol* 2005;**3**:926–9.

171. Gilmore IT, Burroughs A, Murray-Lyon IM, Williams R, Jenkins D, Hopkins A. Indications, methods, and outcomes of percutaneous liver biopsy in England and Wales: an audit by the British Society of Gastroenterology and the Royal College of Physicians of London. *Gut* 1995;**36**:437–41.

172. Janes CH, Lindor KD. Outcome of patients hospitalized for complications after outpatient liver biopsy. *Ann Intern Med* 1993;**118**:96–8.

173. Lindor KD, Bru C, Jorgensen RA, Rakela J, Bordas JM, Gross JB, *et al.* The role of ultrasonography and automatic-needle biopsy in outpatient percutaneous liver biopsy. *Hepatology* 1996;**23**:1079–83.

174. Little AF, Ferris JV, Dodd GD. Image-guided percutaneous hepatic biopsy: effect of ascites on the complication rate. *Radiology* 1996;**199**:79–83.

175. Maharaj B, Bhoora IG. Complications associated with percutaneous needle biopsy of the liver when one, two or three specimens are taken. *Postgrad Med J* 1992;**68**:964–7.

176. Manolakopoulos S, Triantos C, Bethanis S, Theodoropoulos J, Vlachogiannakos J, Cholongitas E, *et al.* Ultrasound-guided liver biopsy in real life: comparison of same-day prebiopsy versus real-time ultrasound approach. *J Gastroenterol Hepatol* 2007;**22**:1490–3.

177. McVay PA, Toy PT. Lack of increased bleeding after liver biopsy in patients with mild hemostatic abnormalities. *Am J Pathol* 1990;**94**:747–53.

178. Sheets PW, Brumbaugh CJ, Kopecky KK. Safety and efficacy of a spring-propelled 18-gauge needle for US-guided liver biopsy. *J Vase Interv Radiol* 1991;**2**:147–9.

179. Smith BC, Desmond PV. Outpatient liver biopsy using ultrasound guidance and the biopty gun is safe and cost effective. *Aust New Zeal J Med* 1995;**25**:209–11.

180. Thampanitchawong P, Piratvisuth T. Liver biopsy: complications and risk factors. *World J Gastroenterol* 1999;**5**:301–4.

181. van der Poorten D, Kwok A, Lam T, Ridley L, Jones DB, Ngu MC, *et al.* Twenty-year audit of percutaenous liver biopsy in a major Australian teaching hospital. *Intern Med J* 2006;**36**:692–9.

182. Vivas S, Palacio MA, Rodriguez M, Lomo J, Cadenas F, Giganto F, *et al.* Biopsia hepática ambulatoria: complicaciones y evolución en 264 casos/ambulatory liver biopsy: complications and evolution in 264 cases. *Rev Esp Enferm Dig* 1998;**90**:175–82.

183. Chevallier P, Ruitort F, Denys A, Staccini P, Saint-Paul MC, Ouzan D, *et al.* Influence of operator experience on performance of ultrasound-guided percutaneous liver biopsy. *Eur Radiol* 2004;**14**:2086–91.

184. Colombo M, Del Ninno E, de Franchis R, De Fazio C, Festorazzi S, Ronchi G, *et al.* Ultrasound-assisted percutaneous liver biopsy: superiority of the Tru-Cut over the Menghini needle for diagnosis of cirrhosis. *Gastroenterology* 1988;**95**:487–9.

185. Denzer U, Arnoldy A, Kanzler S, Galle PR, Dienes HP, Lohse AW. Prospective randomized comparison of minilaparoscopy and percutaneous liver biopsy: diagnosis of cirrhosis and complications. *J Clin Gastroenterol* 2007;**41**:103–10.

186. Farrell RJ, Smiddy PF, Pilkington RM, Tobin AA, Mooney EE, Temperley IJ, *et al.* Guided versus blind liver biopsy for chronic hepatitis C: clinical benefits and costs. *J Hepatol* 1999;**30**:580–7.

187. Papini E, Pacella CM, Rossi Z, Bizzarri G, Fabbrini R, Nardi F, *et al.* A randomized trial of ultrasound-guided anterior subcostal liver biopsy versus the conventional Menghini technique. *J Hepatol* 1991;**13**:291–7.

188. Terjung B, Lemnitzer I, Dumoulin FL. Bleeding complications after percutaneous liver biopsy. An analysis of risk factors. *Digestion* 2003;**67**:138–45.

189. Wawrzynowicz-Syczewska M, Kruszewski T, Boron-Kaczmarska A. Complications of percutaneous liver biopsy. *Rom J Gastroenterol* 2002;**11**:105–7.

190. Weigand K. Percutaneous liver biopsy: retrospective study over 15 years comparing 287 inpatients with 428 outpatients. *J Gastroenterol Hepatol* 2009;**24**:792–9.

191. Choh J, Dolmatch B, Safadi R, Long P, Geisinger M, Lammert G, *et al.* Transjugular core liver biopsy with a 19-gauge spring-loaded cutting needle. *Cardiovasc Intervent Radiol* 1998;**21**:88–90.

192. Papatheodoridis GV, Patch D, Watkinson A, Tibballs J, Burroughs AK. Transjugular liver biopsy in the 1990s: a 2-year audit. *Aliment Pharmacol Ther* 1999;**13**:603–8.

193. Soyer P, Fargeaudou Y, Boudiaf M, Rymer R. Transjugular liver biopsy using ultrasonographic guidance for jugular vein puncture and an automated device for hepatic tissue sampling: a retrospective analysis of 200 consecutive cases. *Abdom Imaging* 2008;**33**:627–32.

194. Vlavianos P, Bird G, Portmann B, Westaby D, Williams R. Transjugular liver biopsy: Use in a selected high risk population. *Eur J Gastroenterol Hepatol* 1991;**3**:469–472.

195. Shaheen AA, Wan AF, Myers RP. FibroTest and FibroScan for the prediction of hepatitis C-related fibrosis: a systematic review of diagnostic test accuracy. *Am J Gastroenterol* 2007;**102**:2589–600.

196. Calès P, Oberti F, Rousselet MC, Gallois Y, Bedossa P. Usual biochemical tests are not suitable surrogate markers for the degree of liver fibrosis. *J Hepatol* 2002;**36**:60.

197. Calès P, Oberti F, Michalak S, Hubert-Fouchard I, Rousselet MC, Konate A, *et al.* A novel panel of blood markers to assess the degree of liver fibrosis. *Hepatology* 2005;**42**:1373–81.

198. Clevert D, Zachoval R, Reiser M. Acoustic radiation force impulse technology and Fibroscan in the evaluation of liver diseases. *AJR Am J Roentgenol* 2009;**192**(Suppl. 5):A72.

199. de Ledinghen V, Laharie D, Foucher J, Bernard P, Castera L, Adhoute X, *et al.* Assessment of cirrhosis and its severity by Fibroscan(R) and biochemical markers in alcoholic patients. 2006. Poster presentation at Digestive Disease Week, Los Angeles, 2006.

200. Foucher J, Castera L, Bernard PH, Adhoute X, Bertet J, Couzigou P, *et al.* Assessment of cirrhosis and its severity by Fibroscan and biochemical markers in alcoholic patients. *J Hepatol* 2006;**44**:S39.

201. Foucher J, Chanteloup E, Vergniol J, Castéra L, Le Bail B, Adhoute X, *et al.* Diagnosis of cirrhosis by transient elastography (FibroScan): a prospective study. *Gut* 2006;**55**:403–8.

202. Jimenez-Ridruejo JM, Gomez-Dominguez E, Otero RM. Measurement of hepatic elasticity: potential usefulness of Fibroscan. *Gastroenterologia y Hepatologia Continuada* 2008;**7**:27–30.

203. Laharie D, Foucher J, Bernard PH, Castera L, Adhoute X, Bertet J, *et al.* Assessment of cirrhosis and its severity by Fibroscan and biochemical markers in alcoholic patients. *Gastroenterology* 2006;**130**:A805.

204. Lee da M, Moon EJ, Hwang JA, Lee MS, Cheong JY, Cho SW, *et al.* Factors associated with liver stiffness in chronic liver disease. *Korean J Hepatol* 2009;**15**:464–73.

205. Marin-Gabriel JC, De la Cruz JB, Tocado M, Rodriguez-Gil Y, Fernandez-Vazquez I, Martin-Algibez AM, *et al.* Failure of liver stiffness measurement with Fibroscan: prevalence and prognostic score. *Gastroenterology* 2008;**134**:A827.

206. Melin P, Dacon A, Gauchet A, Schoeny M, Diebold MD. Interest of the FibroScan in the screening of cirrhosis in patients attending alcoholism consultation. *J Hepatol* 2005;**42**:685.

207. Morozov SV, Trufanova, Iu M, Pavlova TV, Isakov VA, Kaganov BS. Elastography for determination of hepatic fibrosis severity: registration study in Russia. *Eksp Klin Gastroenterol* 2008;**2**:40–7.

208. Mueller S, Millonig G, Sarovska L, Friedrich S, Reimann F, Pritsch M, *et al.* Alcoholic steatohepatitis increases liver stiffness independent of fibrosis stage: criteria for noninvasive fibrosis assessment. *J Hepatol* 2009;**50**:S366

209. Naveau S, Raynard B, Ratziu V, Abella A, Imbert-Bismut F, Messous D, *et al.* Diagnostic value of biochemical markers (FibroTest) for the prediction of liver fibrosis in patients with chronic alcoholic liver disease (ALD). *Hepatology* 2003;**38**:673A.

210. Naveau S, Gaudé G, Asnacios A, Agostini H, Abella A, Barri-Ova N, *et al.* Diagnostic and 10-year prognostic values of non-invasive biomarkers of fibrosis in patients with Alcoholic liver disease (ALD). *Hepatology* 2008;**48**:637A.

211. Nguyen-Khac E, Robert B, Brevet M, Tramier B, Decrombecque C, Joly JP, *et al.* Assessment of asymptomatic liver fibrosis in alcohol abuser patients by transient elastography (Fibroscan). *J Hepatol* 2007;**46**:S280.

212. Parkes J, Roderick P, Harris S, Gough C, Wheatley M, Alexander GJ, *et al.* European liver fibrosis (ELF) panel of serum markers can predict clinical outcome in a cohort of patients from England with mixed aetiology chronic liver disease. *Hepatology* 2007;**46**:832A.

213. Rosenberg WM, Burt A, Hubscher S, Roskams T, Voelker M, Becka M, *et al.* Serum markers predict liver fibrosis. *Hepatology* 2001;**34**:396A.

214. Thabut D, Lebrec D, Imbert-Bismut F, Cazals-Hatem D, Moreau R, Messous D, *et al.* Diagnostic value of fibrosis biochemical markers (FibroTest) for the prediction of portal hypertension in chronic liver disease. *Hepatology* 2003;**38**:292A.

215. Galena HJ. Complications occurring from diagnostic venipuncture. *J Fam Pract* 1992;**34**:582–4.

216. Horowitz SH. Peripheral nerve injury and causalgia secondary to routine venipuncture. *Neurology* 1994;**44**:964.

217. Stitik TP, Foye PM, Nadler SF, Brachman GO. Phlebotomy-related lateral antebrachial cutaneous nerve injury. *Am J Phys Med Rehabil* 2001;**80**:230–4.

218. Deacon B, Abramowitz J. Fear of needles and vasovagal reactions among phlebotomy patients. *J Anxiety Disord* 2006;**20**:946–60.

219. Godwin PG, Cuthbert AC, Choyce A. Reducing bruising after venepuncture. *Qual Health Care* 1992;**1**:245–6.

220. Choffel C, Duterque M, Milleron B, Bernard A. Sarcoid nodules of the skin at the site of venipuncture. *Nouv Presse Med* 1981;**10**:697–9.

221. Nouri M, Rozema C, Nouri M, Rouchet M, Bailly M. Radial neuropathy after peripheral venous puncture. *Ann Fr Anesth Reanim* 2000;**19**:39–41.

222. Pradhan S, Gupta A. Iatrogenic median and femoral neuropathy. *J Assoc Physicians India* 1995;**43**:141.

223. Saeed MA, Gatens PF. Anterior interosseous nerve syndrome: unusual etiologies. *Arch Phys Med Rehabil* 1983;**64**:182.

224. Sander HW, Conigliari MF, Masdeu JC. Antecubital phlebotomy complicated by lateral antebrachial cutaneous neuropathy. *N Engl J Med* 1998;**339**:2024.

225. Vidal D, Barnadas M, Perez M, Coll P, Alomar A. Tuberculous gumma following venepuncture. *Br J Dermatol* 2001;**144**:601–3.

226. Zubairy AI. How safe is blood sampling? Anterior interosseus nerve injury by venepuncture. *Postgrad Med J* 2002;**78**:625.

227. Rodriguez Guerrero FJ, Seda Diestro J, Martin Llamas J. Aparición de hematomas asociados a la extracción de sangre venosa mediante vacío. *Enferm Clin* 2003;**13**:81–6.

228. Berry PR, Wallis WE. Venepuncture nerve injuries. *Lancet* 1977;**1**:1236–7.

229. Burgdorf WH, Hoxtell EO, Bart BJ. Sarcoid granulomas in venipuncture sites. *Cutis* 1979;**24**:52–3.

230. Norcross WA, Shackford SR. Arteriovenous fistula. A potential complication of venipuncture. *Arch Intern Med* 1988;**148**:1815–6.

231. Yuan RT, Cohen MJ. Lateral antebrachial cutaneous nerve injury as a complication of phlebotomy. *Plast Reconstr Surg* 1985;**76**:299–300.

232. Wakita R, Ohno Y, Yamazaki S, Kohase H, Umino M. Vasovagal syncope with asystole associated with intravenous access. *Oral Surge Oral Med Oral Pathol Oral Radiol Endod* 2006;**102**:e28–e32.

233. Horowitz SH. Venipuncture-induced causalgia: anatomic relations of upper extremity superficial veins and nerves, and clinical considerations. *Transfusion* 2000;**40**:1036–40.

234. Newman BH, Waxman DA. Blood donation-related neurologic needle injury: evaluation of 2 years' worth of data from a large blood center. *Transfusion* 1996;**36**:213–15.

# Appendix 1

## Categorisation of disease progression as identified by liver biopsy

In chronic liver disease, the liver may be affected by inflammation or fibrosis or both. The term 'grading' is conventionally used to describe the degree of inflammatory activity, whereas the term 'staging' is used to describe the degree of fibrosis and also architectural change.[15]

A number of systems have been developed to measure and categorise these factors as identified by liver biopsy. These include the Knodell Histological Activity Index (HAI),[164] the Ishak-modified HAI,[165] and the Scheuer,[166] Batts–Ludwig,[134] Brunt,[137] and METAVIR[133] scoring systems. Some of these were developed for chronic liver disease of a specific aetiology, for example Brunt[137] and Kleiner[136] for NAFLD and METAVIR for hepatitis C.[94]

The studies reviewed in this report that evaluated the test accuracy of NILTs using liver biopsy as their reference standard most commonly used the METAVIR staging system to measure the degree of fibrosis; however, some used other systems. For comparative purposes, the staging systems used in the included studies are summarised in *Table 36*. It should be noted that, although the staging systems have numeric labels that imply a linear increase in fibrosis severity between stages, they are in fact ordinal. In other words, although the stages follow a logical ordering in terms of disease severity, they do not represent an underlying continuous scale of measurement such that equal differences between values in the scale represent equivalent differences in the degree of fibrosis (so, for instance, METAVIR stage F4 does not indicate twice as much fibrosis as METAVIR stage F2). However, the degree of fibrosis is a continuous variable and the non-invasive tests reviewed in this report measure continuous variables (serum biochemical markers or liver stiffness) yielding results that may occur at any point on the relevant scale.

**TABLE 36** Fibrosis staging systems used in included studies

| Stage/degree of fibrosis | | | | | | |
|---|---|---|---|---|---|---|
| **METAVIR[94]** | **Scheuer 1991[166]** | **Brunt 1999[137]** | **Kleiner 2005[136]** | **Batts–Ludwig 2010[134]** | **Knodell 1981[164]** | **Ishak modified HAI[165]** |
| 0: No fibrosis | 0: No fibrosis | 0: No fibrosis | 0: No fibrosis | 0: No fibrosis | 0: No fibrosis | 0: No fibrosis |
| 1: Stellate enlargement of portal tract, but without septa formation | 1: Enlarged fibrotic portal tracts | 1: Zone 3 perisinusoidal/ pericellular fibrosis; focally or extensively present | 1: Perisinusoidal or periportal<br>1A: Mild, zone 3, perisinusoidal<br>1B: Moderate, zone 3, perisinusoidal<br>1C: Portal/periportal | 1: Portal fibrosis (fibrous portal expansion) | 1: Fibrous portal expansion | 1: Fibrous expansion of some portal areas, with or without short fibrous septa<br>2: Fibrous expansion of most portal areas, with or without short fibrous septa |
| 2: Enlargement of portal tract with rare septa formation | 2: Periportal or portal–portal septa, but intact architecture | 2: Zone 3 perisinusoidal/ pericellular fibrosis with focal or extensive periportal fibrosis | 2: Perisinusoidal and portal/ periportal | 2: Periportal fibrosis (periportal or rare P–P septa) | 2: Not allocated | 3: Fibrous expansion of some portal areas, with occasional P–P bridging |
| 3: Numerous septa without fibrosis | 3: Fibrosis with architectural distortion, but no obvious cirrhosis | 3: Zone 3 perisinusoidal/ pericellular fibrosis and portal fibrosis with focal or extensive bridging fibrosis | 3: Bridging fibrosis | 3: Septal fibrosis (fibrous septa with architectural distortion, but no obvious cirrhosis) | 3: Bridging fibrosis (P–P or P–C linkage) | 4: Fibrous expansion of some portal areas, with marked bridging (P–P as well as P–C) |
| 4: Cirrhosis | 4: Probable or definite cirrhosis | 4: Cirrhosis | 4: Cirrhosis | 4: Cirrhosis | 4: Cirrhosis | 5: Marked bridging (P–P and/or P–C) with occasional nodules (incomplete cirrhosis)<br>6: Cirrhosis, probable or definite |

P–C, portal to central; P–P, portal to portal.

# Appendix 2

## The Alcohol Use Disorders Identification Test

The AUDIT provides a score based on the following series of questions.[31]

1. How often do you have a drink containing alcohol?

   Never (0)
   Monthly or less (1)
   Two to four times a month (2)
   Two to three times a week (3)
   Four or more times a week (4)

2. How many drinks containing alcohol do you have in a typical day when you are drinking?

   1 or 2 (0)
   3 or 4 (1)
   5 or 6 (2)
   7 to 9 (3)
   10 or more (4)

3. How often do you have six or more drinks on any one occasion?

   Never (0)
   Less than monthly (1)
   Monthly (2)
   Weekly (3)
   Daily or almost daily (4)

4. How often during the last year have you found that you were not able to stop drinking once you had started?

   Never (0)
   Less than monthly (1)
   Monthly (2)
   Weekly (3)
   Daily or almost daily (4)

5. How often during the last year have you failed to do what was normally expected of you because of drinking?

   Never (0)
   Less than monthly (1)
   Monthly (2)
   Weekly (3)
   Daily or almost daily (4)

6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?

   Never (0)
   Less than monthly (1)
   Monthly (2)
   Weekly (3)
   Daily or almost daily (4)

7. How often during the last year have you had a feeling of guilt or remorse after drinking?

   Never (0)
   Less than monthly (1)
   Monthly (2)
   Weekly (3)
   Daily or almost daily (4)

8. How often during the last year have you been unable to remember what happened the night before because you had been drinking?

   Never (0)
   Less than monthly (1)
   Monthly (2)
   Weekly (3)
   Daily or almost daily (4)

9. Have you or someone else been injured because of your drinking?

   No (0)
   Yes, but not in the last year (2)
   Yes, during the last year (4)

10. Has a relative, friend, doctor or other health worker been concerned about your drinking or suggested that you should cut down?

    No (0)
    Yes, but not in the last year (2)
    Yes, during the last year (4)

# Appendix 3

## Liver biopsy: systematic review of adverse events

A systematic review was carried out with the aim of identifying studies that reported adverse events in adults undergoing either percutaneous or transjugular liver biopsy for any form of suspected liver disease. Because the aim was to assess the incidence of adverse events relating to each biopsy route, studies that included < 100 relevant patients undergoing biopsy by each route were excluded, as were any studies in which results were not provided separately for each route.

Details of the literature searches and the inclusion and exclusion criteria may be found in *Appendix 4*.

The searches were not restricted by study type, language, or date. However, because of the possibility of changes in standards of care over time, studies were excluded if they included data relating to biopsies undertaken prior to 1980. Moreover, because of time constraints, it was not possible to include non-English-language papers in the systematic review.

The same three-stage sifting process was used as in the systematic review of clinical effectiveness. Data were extracted directly to the tables included in the report.

The electronic literature searches identified 2289 potentially relevant articles. Of these, 12 met the review's inclusion criteria (for PRISMA diagram, see *Appendix 4*). Seventeen additional relevant articles were identified from citations.[96,167–182]

Data from these studies relating to adverse events, both minor and severe, and deaths associated with liver biopsy are set out in *Tables 37* and *38*. Because patients undergoing transjugular biopsy are at higher risk of complications than those undergoing percutaneous biopsy, the two techniques have been considered separately.

The pooled data for percutaneous biopsy suggest that over 7% of patients suffer related minor adverse events, although < 1% suffer severe related adverse events (including death) and < 0.1% die (see *Table 37*). The comparable figures for transjugular biopsy are somewhat higher, at over 9%, 1.45%, and 0.18%, respectively (see *Table 38*). However, as may be seen, rates vary between individual studies. This is due, at least in part to, the use of varying definitions of minor and severe adverse events. Because of time constraints, it has not been possible to explore these issues, nor those relating to the use of ultrasound guidance for percutaneous biopsy, the clinical setting, and patient and clinician characteristics.

**TABLE 37** Uncontrolled observations of adverse events and mortality associated with liver biopsy: percutaneous biopsy

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Beddy 2007[167] | Platelet count >50,000/mm$^3$, prothrombin time prolonged by <3 seconds, without severe ascites or intrahepatic biliary dilatation | 500 | Ultrasound-guided, outpatient, radiologist | Prospective<br>Multinational<br>Not clear; presumed from publication date to post-date 1980 | Pain requiring analgesia | 15 | 3.00 *(1.64 to 4.53)* | Hospital admission for haemorrhagic complications | 1 | 0.20 (*0.02 to 0.93)* | 0 | 0.00 (0.00 to 0.50) |
| Caturelli 1996[168] | Suspected chronic liver disease, platelet count >40,000/mm$^3$ and/or prothrombin activity <40% of control value | 753 | Ultrasound used to identify biopsy site, 88% outpatients, three skilled gastro-enterologists | Retrospective audit<br>Italy<br>Not clear; presumed from publication date to post-date 1980[a] | Pain or haematoma at biopsy site, shoulder pain | *33* | *4.40 (2.91 to 5.71)* | Haemoperitoneum, vasovagal reaction requiring atropine | 4 | *0.53 (0.18 to 1.25)* | 0 | 0.00 (0.00 to 0.33) |
| Chevallier 2004[183] | Patients mainly had HCV or ALD; exclusions included ascites and clotting disorders | 732 | Ultrasound used to identify biopsy site, some outpatients, experienced and inexperienced radiologists | Comparative (experienced vs inexperienced operators)<br>France<br>1998–2000 | Minor vasovagal reaction<br>Significant pain 6 hours after biopsy | 7<br>*44* | *0.96 (0.38 to 1.76)*<br>6.00 *(4.16 to 7.40)* | Severe vasovagal reaction, haemorrhage | 2 | *0.27 (0.06 to 0.87)* | 0 | *0.00 (0.00 to 0.34)* |

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) | |
| Colombo 1988[184] | Diffuse liver disease, prothrombin and partial thromboplastin times within 3 SD of the normal mean, platelet count > 80,000/mm$^3$, bleeding times < 7 minutes | 1192 | Ultrasound used to identify biopsy site, inpatient, four equally experienced physicians | Prospective (RCT comparing two needles) Italy 1982–5 | Significant pain requiring analgesia, > 5% haematocrit drop not requiring blood transfusion, mild transient vagal shock | *66* | *5.54 (4.08 to 6.54)* | Acute abdomen (resolved in 24 hours without treatment) | *1* | *0.08 (0.01 to 0.39)* | 0 | *0.00 (0.00 to 0.21)* | |
| Denzer 2007[185] | Adults with chronic liver disease or liver disease of unknown aetiology, without severe coagulation defects | 415 | Ultrasound used to identify biopsy site, inpatient and outpatient, four experienced investigators | Prospective (one arm of RCT) Germany NR | Pain, vasovagal symptoms, agitation/ disquiet, minor intra-abdominal bleeding | 19 | *4.58 (2.66 to 6.49)* | Complications requiring hospital stay > 24 hours or new hospitalisation (intra-abdominal bleeding or haemobilia) | 4 | *0.96 (0.32 to 2.25)* | 0 | *0.00 (0.00 to 0.60)* | |
| Douds 1995[169] | Not specified | 453 | Ultrasound used to identify biopsy site, day case or inpatient, no details regarding operator | Retrospective audit UK 1989–93 | Abdominal pain | 7 | *1.55 (0.62 to 2.82)* | Haemorrhage, pneumothorax | 5 | *1.10 (0.36 to 2.23)* | 1 | *0.22 (0.02 to 1.02)* | |

*continued*

**TABLE 37** Uncontrolled observations of adverse events and mortality associated with liver biopsy: percutaneous biopsy *(continued)*

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Farrell 1999[186] | Patients with chronic hepatitis C; platelet count>100,000/mm$^3$, prothrombin time prolonged by <2 seconds | 166 | Ultrasound guidance used in 53% of cases, apparently inpatient, blind biopsies performed by experienced gastroenterology registrar, guided biopsies by experienced radiology registrar or consultant radiologist | Prospective comparison of blind vs guided biopsy<br>Ireland<br>1993–5 | Pain-associated vasovagal symptoms, persistent biopsy-site pain | *9* | *5.42 (2.39 to 8.86)* | Not given | 0 | 0.00 *(0.00 to 1.50)* | 0 | *0.00 (0.00 to 1.50)* |
| Firpi 2005[170] | Any patients undergoing biopsy | 3214 | Ultrasound guidance used to identify biopsy site from February 2002, outpatient, no information regarding operator | Retrospective audit<br>USA<br>1995–2004 | Mild/moderate pain requiring limited analgesia<br>Minor bleeding at biopsy site<br>Nausea and vomiting | *418*<br>*129*<br>*16* | 13 *(11.84 to 14.16)*<br>4.00 (3.32 *to* 4.68)<br>0.50 *(0.26 to 0.74)* | Events resulting in death or hospitalisation either immediately or up to 2 weeks after biopsy | 34 | *1.06 (0.73 to 1.42)* | 2 | *0.06 (0.01 to 0.68)* |
| Gilmore 1995[171] | Any patients undergoing biopsy | 1500 | Pre-biopsy imaging (usually ultrasound) used in 84%; 96% inpatient; 34% of operators radiologists, 33% general internal medicine, 28% gastroenterology, 5% other | Retrospective audit<br>UK<br>1991 | Minor complications NR | NR | | Bleeding | 26 | *1.73 (1.12 to 2.41)* | 5 | 0.33 *(0.11 to 0.68)* |

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Janes 1993[172] | Patients with chronic liver diseases | 405 | Apparently no ultrasound guidance, outpatient, gastro-enterologists (junior doctors or staff physician) | Retrospective<br>USA<br>1989–91 | Minor complications NR | NR | | Hospital admission | 13 | *3.21 (1.67 to 4.97)* | 0 | *0.00 (0.00 to 0.62)* |
| Lindor 1996[173] | Excluded patients with liver transplant, history of previous biopsy complications or difficulty obtaining tissue, suspected liver mass or known haemangiomas in right lobe, known bleeding tendencies or significant ascites | 836 | With or without ultrasound guidance, outpatient, consultants, trained fellows, and experienced associates | Prospective, controlled<br>USA/Spain<br>1992–4 | Bleeding or hypotension not requiring hospitalisation | *16* | *1.91 (1.08 to 2.89)* | Hospitalisation for pain with or without hypotension, lacerated gall bladder | 11 | *1.32 (0.65 to 2.16)* | 0 | *0.00 (0.00 to 0.30)* |
| Little 1996[174] | Patients with suspected focal hepatic lesions, diffuse liver disease, or malignant portal vein thrombus | 476 | Ultrasound or computerised tomography guidance, inpatient and outpatient, no details regarding operators | Retrospective audit<br>USA<br>1992–4 | Decrease in haemoglobin or haematocrit value not requiring treatment | 25 | *5.25 (3.26 to 7.06)* | Bleeding necessitating blood transfusion or surgery, or resulting in death | 16 | *3.36 (1.87 to 4.99)* | 0 | *0.00 (0.00 to 0.53)* |
| McVay 1990[177] | Patients with mild haemostatic abnormalities | 177 | Apparently no ultrasound guidance, inpatient, gastroenterology fellows, radiology residents or attending physicians | Retrospective audit<br>USA<br>1983–7 | Bleeding complications not requiring transfusions | 2 | *1.13 (023 to 3.54)* | Bleeding complications requiring transfusions | 6 | *3.39 (1.22 to 6.29)* | 1 | *0.56 (0.06 to 2.60)* |

*continued*

**TABLE 37** Uncontrolled observations of adverse events and mortality associated with liver biopsy: percutaneous biopsy *(continued)*

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Maharaj 1992[175] | Any patients undergoing biopsy | 2646 | Apparently no ultrasound guidance, inpatient, medical and surgical interns | Retrospective (1984–7) Prospective (1988–90) South Africa | Pain responding to analgesics | 31 | *1.17 (0.79–1.60)* | Intraperitoneal bleeding, pain, symptomatic hypotension, biliary peritonitis | 32 | *1.21 (0.82 to 1.64)* | 8 | *0.30 (0.13 to 0.54)* |
| Manolakopoulos 2007[176] | Patients with diffuse liver disease | 631 | Pre-biopsy or real-time ultrasound, inpatient, operator experienced gastro-enterologist for pre-biopsy ultrasound, radiologist for real-time ultrasound | Prospective comparison (pre-biopsy vs real-time ultrasound) Greece 1995–2005 | Minor complications not reported | NR | | Need for blood transfusion, surgical intervention, or hospital stay >48 hours post-biopsy | 1 | *0.16 (0.02 to 0.74)* | 0 | *0.00 (0.00 to 0.40)* |
| Papini 1991[187] | Women with suspected chronic liver disease (bleeding time <4 minutes, platelet count >50,000/mm$^3$, prothrombin activity >50%) | 340 | Blind intercostal or ultrasound-guided subcostal, 10% outpatient, three experienced hepatologists | Prospective Italy 1985–9 | Asymptomatic haematoma, hypotension, haemothorax, ileus, self-limiting bleeding in thoracic cavity | 8[b] | *2.35 (1.00 to 4.11)* | Strong pain needing treatment | 10[c] | *2.94 (1.38 to 4.84)* | 0 | *0.00 (0.00 to 0.40)* |
| Sheets 1991[178] | Patients with suspected diffuse liver disease | 114 (203 biopsies) | Ultrasound guidance used, no data regarding inpatient/outpatient status or operator | Not clear USA 1988–9 | Complications not requiring treatment (syncope, severe pain) | 2 | *1.75 (0.36 to 5.42)* | Complications requiring treatment (syncope, bleeding requiring transfusion) | 4 | *3.51 (1.15 to 7.86)* | 0 | *0.00 (0.00 to 2.17)* |
| Smith 1995[179] | Any patients | 250 | Ultrasound guidance used, 70% outpatient, gastro-enterology registrar | Retrospective audit Australia 1991–3 | Prolonged pain requiring narcotic administration without hospital admission | 1 | *0.40 (0.04 to 1.85)* | Prolonged pain requiring narcotic administration and unplanned hospital admission | 3 | *1.20 (0.34 to 3.13)* | 0 | *0.00 (0.00 to 1.00)* |

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Terjung 2003[188] | Any patients undergoing biopsy (40% had risk factors for biopsy-related bleeding) | 574 | Ultrasound used to identify biopsy site, inpatient, seven experienced hepatogastro-enterologists or five hepatogastro-enterology fellows under their supervision | Retrospective audit<br>Germany<br>1993–6 | Fall in serum haemoglobin >2 g/dl,haematoma assessed by ultrasound | 62 | *10.80 (7.57 to 12.17)* | Clinically overt bleeding complication | 10 | *1.74 (0.83 to 2.91)* | 3 | *0.52 (0.15 to 1.38)* |
| Thampanitchawong 1999[180] | Any patients | 484 biopsies | No data regarding use of ultrasound or inpatient/outpatient status, medical staff or residents (after 1994, all were well-trained gastroenterologists or residents under their supervision) | Retrospective audit<br>Thailand<br>1987–96 | Minor bleeding (decrease in haematocrit ≥4% not requiring transfusion or surgery), transient hypotension | 18 | *3.72 (2.14 to 5.38)* | Major bleeding (decrease in haematocrit ≥4% requiring transfusion or surgery) | 17 | *3.51 (1.99 to 5.14)* | 6 | 1.20 *(0.45 to 2.37)* |
| Van der Poorten 2006[181] | Any patients | 1398 | Ultrasound guidance used in 87.9%, 66% outpatient, mostly performed by radiologist or radiology registrar | Retrospective audit<br>Australia<br>1996–2005 | Pain, nausea, vomiting, vasovagal episodes | 166 | *11.87 (9.14 to 12.19)* | Significant haemorrhage, fall in haemoglobin of >20 g/l, bile peritonitis, visceral perforation, death | 12 | *0.86 (0.44 to 1.39)* | 3 | *0.21 (0.06 to 0.57)* |
| Van Thiel 1993[96] | Patients attending liver transplant centre | 12,750 biopsies | Ultrasound guidance used in 0.4%, 95% inpatient, 67% performed by surgeons,33% by transplant physicians, 0.4% by radiologists | Retrospective audit<br>USA<br>1989–91 | Minor complications NR | NR | – | Major complications including intra-abdominal haemorrhage | 26 | *0.20 (0.13 to 0.29)* | 0 | *0.00 (0.00 to 0.02)* |

*continued*

**TABLE 37** Uncontrolled observations of adverse events and mortality associated with liver biopsy: percutaneous biopsy *(continued)*

| Study | Brief description of patient group | Number of patients | Details of biopsy, setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Vivas 1998[182] | Patients with platelet count > 60,000/mm$^3$, prothrombin time < 4 seconds above control, without ascites, HE or liver tumours | 378 | Ultrasound used to identify biopsy site, 70% outpatient, medical staff and residents in their third or fourth year of training | Prospective<br>Spain<br>1995–6 | Pain requiring analgesia, but not hospital admission | 36 | *9.52 (6.18 to 11.59)* | Pain or hypertension requiring hospital admission, abdominal bleeding requiring blood transfusion, subcapsular haematoma | 7 | *1.85 (0.74 to 3.37)* | 0 | *0.00 (0.00 to 0.66)* |
| Wawrzynowicz-Syczewska 2002[189] | Patients with platelet count ≥ 50,000/mm$^3$ and prothrombin activity ≥ 50% | 861 | Ultrasound only occasionally used to identify biopsy site, inpatient, no data regarding operator | Retrospective audit<br>Poland<br>1997–2001 | Severe pain requiring intravenous analgesia or morphine | 60 | *6.97 (5.01 to 8.19)* | Severe vasovagal reaction requiring atropine, haemoperitoneum, pneumothorax, renal puncture, haemothorax, septic shock | 12 | *1.39 (0.71 to 2.25)* | 0 | *0.00 (0.00–0.29)* |
| Weigand 2009[190] | Patients without clotting disorders (prothrombin time < 70%, platelets < 80,000/ml) or anaemia(haemoglobin < 10 g/dl) | 715 | Ultrasound used, 60% outpatient, experienced investigators | Retrospective audit<br>Germany<br>1990–2005 | Minor complications, most commonly mild pain | 63 | *8.81 (6.29 to 10.11)* | Any event resulting in death, severe pain not resolving after oral analgesic, prolonged severe hypotension, or bleeding plus significant drop in haemoglobin level | 2 | *0.28 (0.06 to 0.89)* | 0 | *0.00 (0.00 to 0.35)* |
| Total | | *31,960* | | | | *1253/ 16,674* | *7.51 (7.11 to 7.91)* | | *259/ 31960* | *0.81 (0.71 to 0.91)* | *29/ 31960* | *0.09 (0.06 to 0.12)* |

HCV, hepatitis C virus; NR, not reported.

a Control group (patients in the same hospital undergoing biopsy without ultrasound localisation) excluded because of greater possibility that biopsy dates might precede 1980.

b 7/170 patients having Menghini biopsy, 1/170 having ultrasound-guided anterior subcostal biopsy.

c 6/170 patients having Menghini biopsy, 4/170 having ultrasound-guided anterior subcostal biopsy.

Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

DOI: 10.3310/hta16040





**TABLE 38** Uncontrolled observations of adverse events and mortality associated with liver biopsy: transjugular biopsy

| Study | Brief description of patient group | Number of patients | Details of setting, operator, etc. | Design/study country/date of biopsies | Minor adverse events related to biopsy | | | Severe adverse events (including deaths) related to biopsy | | | Deaths related to biopsy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | Definition | *n* | % (95% CI) | Definition | *n* | % (95% CI) | *n* | % (95% CI) |
| Choh 1998[191] | Patients with coagulopathy and/or ascites | 101 | No data regarding ultrasound guidance, inpatient status or operator | Case series<br>USA<br>1994–6 | Minor neck haematoma | 2 | *1.98 (0.41 to 6.08)* | Not given | 3 | *0.99 (0.82 to 7.50)* | 0 | *0.00 (0.00 to 0.45)* |
| Papatheodoridis 1999[192] | Patients with acute liver disease, chronic liver disease or post-transplant disorder | 145 | No data regarding ultrasound guidance, inpatient status or operator | Prospective<br>UK<br>1995–7 | Minor complications not reported | NR | | Capsular perforations | 2 | *1.38 (0.28 to 4.29)* | 0 | *0.00 (0.00 to 1.71)* |
| Soyer 2008[193] | Any patients undergoing biopsy | 200 | Ultrasound guidance, inpatient, experienced radiologist aided by a technician | Retrospective audit<br>France<br>1995–2007 | Pain at biopsy site or puncture site; mild haematoma at puncture site | 24 | 12.00 *(7.02 to 15.07)* | Not given | 0 | *0.00 (0.00 to 1.25)* | 0 | *0.00 (0.00 to 1.25)* |
| Vlavianos 1991[194] | Patients with prothrombin time > 6 seconds above control, platelet count < 40,000/mm$^3$ or tense ascites | 104 | Inpatient status NR, under supervision of an experienced hepatologist | Prospective<br>UK<br>1985–9 | Perforation of liver capsule without clinical sequelae, tachycardia or small haemothorax which resolved spontaneously | 11 | *10.58 (4.92 to 15.54)* | Ruptured subcapsular haematoma, perforation of liver capsule requiring blood transfusion, pneumothorax requiring drainage | 3 | *2.88 (0.80 to 7.29)* | 1 | *0.96 (0.10 to 4.36)* |
| Total | | *550* | | | | *37/405* | *9.14 (6.33 to 11.95)* | | *8/550* | *1.45 (0.45 to 2.45)* | *1/550* | *0.18 (0.00 to 0.53)* |

NR, not reported.
Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

# Appendix 4

# Systematic review of the adverse effects of liver biopsy: search strategies

## Sources searched

The electronic bibliographic databases that were searched are listed in *Appendix 5*. The searches were carried out in February 2010.

## Search strategies

The MEDLINE search strategy may be found in *Appendix 5*.

### *Inclusion criteria*

#### Population

- Adults.

#### Intervention

- Percutaneous or transjugular liver biopsy.

#### Outcomes

- Adverse events probably or possibly caused by the liver biopsy.

#### Setting

- Any country.

#### Study type

- Any study design which presented data relating to over 100 patients.

### *Exclusion criteria*

#### Population

- Fewer than 100 participants (either overall or in either arm of a RCT comparing percutaneous with transjugular biopsy).

#### Intervention

- Laparoscopic liver biopsy.
- Biopsy undertaken prior to 1980.

#### Study type

- Animal models.
- Narrative reviews, editorials, opinions.

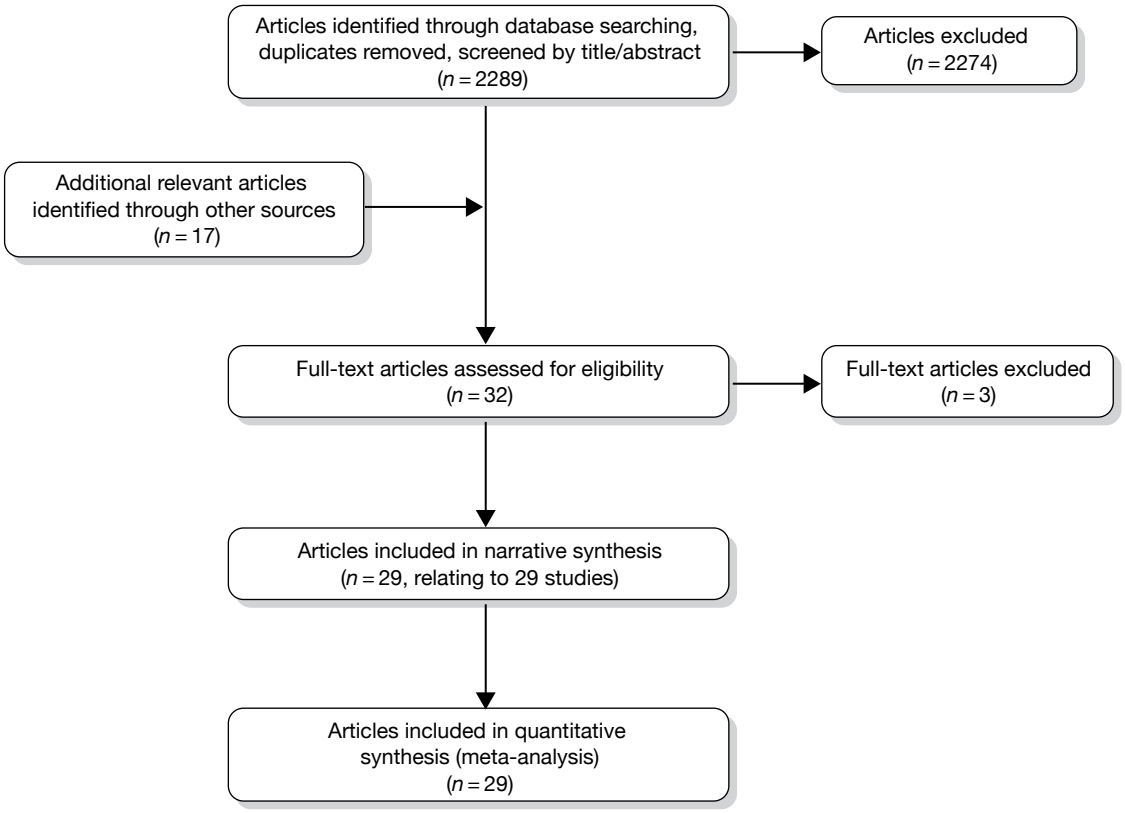



**FIGURE 28** Adverse effects of liver biopsy: summary of study selection and exclusion.

# Appendix 5

# Assessment of clinical effectiveness and cost-effectiveness, adverse effects, and quality of life: search strategies

The following electronic databases were searched:

1. MEDLINE via Ovid (1950 to present)
2. EMBASE via Ovid (1980 to present)
3. MEDLINE In-Process & Other Non-Indexed Citations via Ovid (1950 to present)
4. The Cochrane library:
   - Cochrane Database of Systematic Reviews (CDSR, 1996 to present)
   - Cochrane Central Register of Controlled Trials (CENTRAL, 1898 to present)
   - Database of Abstracts of Reviews of Effects (DARE, 1995 to present)
   - Cochrane Methodology Register (1904 to present)
   - Health Technology Assessment Database (1995 to present)
   - NHS Economic Evaluation Database (1995 to present)
5. CINAHL via EBSCO (1982 to present)
6. Web of Knowledge:
   - Science Citation Index (SCI, 1969 to present)
   - Conference Proceedings Citation Index (CPCI-S, 1990 to present)
   - BIOSIS Previews (1969 to present).

## Search strategies

The search strategies shown below were used in Ovid MEDLINE (1950 to present), and were adapted for use across multiple databases.

### *Clinical effectiveness search strategy*

To retrieve evidence of the diagnostic test reliability and accuracy, the broad and specific intervention terms (1–5, 9–15, 20–60) were combined with those of the clinical condition, namely liver fibrosis (6–7, 17–19), and then combined with the diagnostic filter (62–74). Search terms for the diagnostic test manufacturers were also included in the strategy (75–78).

1. (enhanced adj liver adj fibrosis).tw.
2. (elf adj test$).tw.
3. (elf and diagnos$).tw.
4. (elf and (fibros*s or cirrhos*s)).tw.
5. elf.tw.
6. exp liver cirrhosis/or exp liver diseases, alcoholic/
7. 5 and 6
8. 1 or 2 or 3 or 4 or 7
9. FibroTest.tw.
10. fibrosure.tw.
11. fibromax.tw.
12. FibroScan.tw.

13. ashtest.tw.
14. (transient adj elastograph$).tw.
15. (elastograph$and liver).tw.
16. or/9-15
17. exp liver cirrhosis/or exp liver diseases, alcoholic/
18. (fibros*s or cirrhos*s).tw.
19. 17 or 18
20. Biological Markers/
21. (biomarker$or bio-marker$).tw.
22. (marker$and (biologic$or biochemical or serum or direct or indirect)).tw.
23. Algorithms/
24. algorithm$.tw.
25. (composite and blood).tw.
26. or/20-25
27. 19 and 26
28. Hyaluronic Acid/
29. ((hyaluronic adj acid) or (hyalauronate or hyaluronan)).tw.
30. 28 or 29
31. (procollagen or piiinp or p3np or ppcp).tw.
32. ((tissue and inhibitor and metalloproteinase$) or timps).tw.
33. 30 and 31 and 32
34. 30 or 31 or 32
35. 34 and 19
36. Alpha-Macroglobulins/
37. ((alpha and macroglobulin$) or (alpha adj 2m)).tw.
38. 36 or 37
39. ((apolipoprotein$adj a1) or apoa1).tw.
40. Haptoglobins/
41. haptoglobin$.tw.
42. 40 or 41
43. (bilirubin$or hematoidin$).tw.
44. (gamma adj glutamyl adj transpeptidase$).tw.
45. (gamma adj glutamyltransferase$).tw.
46. ((gamma adj gt) or ggt or ggtp).tw.
47. 44 or 45 or 46
48. 38 and 39 and 42 and 43 and 47
49. 38 or 39 or 42 or 43 or 47
50. 49 and 19
51. (alanine adj (aminotransferase$or aminotransaminase$)).tw.
52. (serum adj glutamic adj pyruvic adj transaminase$).tw.
53. sgpt.tw.
54. 51 or 52 or 53
55. (aspartate adj (aminotransferase$or aminotransaminase$)).tw.
56. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
57. sgot.tw.
58. 55 or 56 or 57
59. 38 and 39 and 42 and 43 and 47 and 54 and 58
60. 38 or 39 or 42 or 43 or 47 or 54 or 58
61. 60 and 19
62. exp “Sensitivity and Specificity”/
63. sensitivity.tw.
64. specificity.tw.

65. ((pre-test or pretest) adj probability).tw.
66. post-test probability.tw.
67. predictive value$.tw.
68. likelihood ratio$.tw.
69. or/62-68
70. 27 and 69
71. 35 and 69
72. 50 and 69
73. 61 and 69
74. 70 or 71 or 72 or 73
75. iqur.tw.
76. biopredictive.tw.
77. echosens.tw.
78. 75 or 76 or 77
79. 8 or 16 or 33 or 48 or 59 or 74 or 78

### *Cost-effectiveness search strategy*

To retrieve evidence of cost-effectiveness, a costs filter (80–100) was added to the search strategy for the clinical effectiveness studies (1–79).

1. (enhanced adj liver adj fibrosis).tw.
2. (elf adj test$).tw.
3. (elf and diagnos$).tw.
4. (elf and (fibros*s or cirrhos*s)).tw.
5. elf.tw.
6. exp liver cirrhosis/or exp liver diseases, alcoholic/
7. 5 and 6
8. 1 or 2 or 3 or 4 or 7
9. FibroTest.tw.
10. fibrosure.tw.
11. fibromax.tw.
12. FibroScan.tw.
13. ashtest.tw.
14. (transient adj elastograph$).tw.
15. (elastograph$and liver).tw.
16. or/9-15
17. exp liver cirrhosis/or exp liver diseases, alcoholic/
18. (fibros*s or cirrhos*s).tw.
19. 17 or 18
20. Biological Markers/
21. (biomarker$or bio-marker$).tw.
22. (marker$and (biologic$or biochemical or serum or direct or indirect)).tw.
23. Algorithms/
24. algorithm$.tw.
25. (composite and blood).tw.
26. or/20-25
27. 19 and 26
28. Hyaluronic Acid/
29. ((hyaluronic adj acid) or (hyalauronate or hyaluronan)).tw.
30. 28 or 29
31. (procollagen or piiinp or p3np or ppcp).tw.
32. ((tissue and inhibitor and metalloproteinase$) or timps).tw.

33. 30 and 31 and 32
34. 30 or 31 or 32
35. 34 and 19
36. Alpha-Macroglobulins/
37. ((alpha and macroglobulin$) or (alpha adj 2m)).tw.
38. 36 or 37
39. ((apolipoprotein$adj a1) or apoa1).tw.
40. Haptoglobins/
41. haptoglobin$.tw.
42. 40 or 41
43. (bilirubin$or hematoidin$).tw.
44. (gamma adj glutamyl adj transpeptidase$).tw.
45. (gamma adj glutamyltransferase$).tw.
46. ((gamma adj gt) or ggt or ggtp).tw.
47. 44 or 45 or 46
48. 38 and 39 and 42 and 43 and 47
49. 38 or 39 or 42 or 43 or 47
50. 49 and 19
51. (alanine adj (aminotransferase$or aminotransaminase$)).tw.
52. (serum adj glutamic adj pyruvic adj transaminase$).tw.
53. sgpt.tw.
54. 51 or 52 or 53
55. (aspartate adj (aminotransferase$or aminotransaminase$)).tw.
56. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
57. sgot.tw.
58. 55 or 56 or 57
59. 38 and 39 and 42 and 43 and 47 and 54 and 58
60. 38 or 39 or 42 or 43 or 47 or 54 or 58
61. 60 and 19
62. exp “Sensitivity and Specificity”/
63. sensitivity.tw.
64. specificity.tw.
65. ((pre-test or pretest) adj probability).tw.
66. post-test probability.tw.
67. predictive value$.tw.
68. likelihood ratio$.tw.
69. or/62-68
70. 27 and 69
71. 35 and 69
72. 50 and 69
73. 61 and 69
74. 70 or 71 or 72 or 73
75. iqur.tw.
76. biopredictive.tw.
77. echosens.tw.
78. 75 or 76 or 77
79. 8 or 16 or 33 or 48 or 59 or 74 or 78
80. exp “Costs and Cost Analysis”/
81. Economics/
82. exp Economics, Hospital/
83. exp Economics, Medical/
84. Economics, Nursing/

85. exp models, economic/
86. Economics, Pharmaceutical/
87. exp "Fees and Charges"/
88. exp Budgets/
89. budget$.tw.
90. ec.fs.
91. cost$.ti.
92. (cost$adj2 (effective$or utilit$or benefit$or minimi$)).ab.
93. (economic$or pharmacoeconomic$or pharmaco-economic$).ti.
94. (price$or pricing$).tw.
95. (financial or finance or finances or financed).tw.
96. (fee or fees).tw.
97. (value adj2 (money or monetary)).tw.
98. quality-adjusted life years/
99. (qaly or qalys).af.
100. (quality adjusted life year or quality adjusted life years).af.
101. or/80-100
102. 79 and 101

## Adverse events searches

### *Venepuncture and transient elastography*

A search strategy was developed to search for the adverse effects of venepuncture and transient elastography. This strategy included both subject headings with adverse effect subheadings for blood tests and imaging techniques (1–5) and free-text terms for adverse effects (7–13) combined with the statements for the diagnostic test interventions (15–25). The search was limited to the adult population (27–28).

1. exp Hematologic Tests/ae [Adverse Effects]
2. exp Serologic Test/ae [Adverse Effects]
3. Blood Specimen Collection/ae [Adverse Effects]
4. Phlebotomy/ae [Adverse Effects]
5. Elasticity Imaging Techniques/ae [Adverse Effects]
6. or/1-5
7. (adverse adj (event$or effect$or outcome$)).ab,ti.
8. risk$.ab,ti.
9. (safe or safety).ab,ti.
10. harm$.ab,ti.
11. complication$.ab,ti.
12. (treatment adj emergent).ab,ti.
13. tolerability.ab,ti.
14. 7 or 8 or 9 or 10 or 11 or 12 or 13
15. (FibroTest or fibrosure or fibromax or ashtest or FibroScan).tw.
16. (transient adj elastograph$).tw.
17. (elastograph$and liver).tw.
18. (enhanced adj liver adj fibrosis).tw.
19. (elf adj test$).tw.
20. (elf and diagnos$).tw.
21. (elf and (fibros*s or cirrhos*s)).tw.
22. elf.tw.
23. exp liver cirrhosis/or exp liver diseases, alcoholic/

24. 22 and 23
25. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 24
26. 14 and 25
27. adult/or aged/or middle aged/or young adult/
28. adult$.tw.
29. 27 or 28
30. 6 or 26
31. 29 and 30

### *Liver biopsy*

The search strategy which was developed to identify studies of the adverse effects of liver biopsy includes subject headings with adverse effect subheadings for biopsy (1–2) combined with the liver fibrosis terms (4–5). Free-text terms for adverse effects (8–14) were combined with the statement liver biopsy (16). The search was limited to the adult population (27–28).

1. Biopsy/ae [Adverse Effects]
2. exp Biopsy, Needle/ae [Adverse Effects]
3. 1 or 2
4. exp liver cirrhosis/or exp liver diseases, alcoholic/
5. (cirrhos*s or fibros*s).tw.
6. 4 or 5
7. 3 and 6
8. (adverse adj (event$or effect$or outcome$)).ab,ti.
9. risk$.ab,ti.
10. (safe or safety).ab,ti.
11. harm$.ab,ti.
12. complication$.ab,ti.
13. (treatment adj emergent).ab,ti.
14. tolerability.ab,ti.
15. 8 or 9 or 10 or 11 or 12 or 13 or 14
16. (liver and biops$).tw.
17. 15 and 16
18. 6 and 17
19. 7 or 18
20. letter.pt.
21. editorial.pt.
22. comment.pt.
23. 20 or 21 or 22
24. 19 not 23
25. adult/or aged/or middle aged/or young adult/
26. adult$.tw.
27. 25 or 26
28. 24 and 27

## Quality of life searches

To search for evidence relating to the impact of the diagnostic tests on the well-being of patients with ALD, a quality-of-life search filter (80–115) was combined with the diagnostic test effectiveness searches (1–79). The quality-of-life filter consists of database subject headings and free-text terms associated with the quality of life, including generic, instrument specific, and methodological terms. In addition, searches were also carried out on the health-related quality of life of patients with the medical condition only (116–121).

1. (enhanced adj liver adj fibrosis).tw.
2. (elf adj test$).tw.
3. (elf and diagnos$).tw.
4. (elf and (fibros*s or cirrhos*s)).tw.
5. elf.tw.
6. exp liver cirrhosis/or exp liver diseases, alcoholic/
7. 5 and 6
8. 1 or 2 or 3 or 4 or 7
9. FibroTest.tw.
10. fibrosure.tw.
11. fibromax.tw.
12. FibroScan.tw.
13. ashtest.tw.
14. (transient adj elastograph$).tw.
15. (elastograph$and liver).tw.
16. or/9-15
17. exp liver cirrhosis/or exp liver diseases, alcoholic/
18. (fibros*s or cirrhos*s).tw.
19. 17 or 18
20. Biological Markers/
21. (biomarker$or bio-marker$).tw.
22. (marker$and (biologic$or biochemical or serum or direct or indirect)).tw.
23. Algorithms/
24. algorithm$.tw.
25. (composite and blood).tw.
26. or/20-25
27. 19 and 26
28. Hyaluronic Acid/
29. ((hyaluronic adj acid) or (hyalauronate or hyaluronan)).tw.
30. 28 or 29
31. (procollagen or piiinp or p3np or ppcp).tw.
32. ((tissue and inhibitor and metalloproteinase$) or timps).tw.
33. 30 and 31 and 32
34. 30 or 31 or 32
35. 34 and 19
36. Alpha-Macroglobulins/
37. ((alpha and macroglobulin$) or (alpha adj 2m)).tw.
38. 36 or 37
39. ((apolipoprotein$adj a1) or apoa1).tw.
40. Haptoglobins/
41. haptoglobin$.tw.
42. 40 or 41
43. (bilirubin$or hematoidin$).tw.
44. (gamma adj glutamyl adj transpeptidase$).tw.
45. (gamma adj glutamyltransferase$).tw.
46. ((gamma adj gt) or ggt or ggtp).tw.
47. 44 or 45 or 46
48. 38 and 39 and 42 and 43 and 47
49. 38 or 39 or 42 or 43 or 47
50. 49 and 19
51. (alanine adj (aminotransferase$or aminotransaminase$)).tw.

52. (serum adj glutamic adj pyruvic adj transaminase$).tw.
53. sgpt.tw.
54. 51 or 52 or 53
55. (aspartate adj (aminotransferase$or aminotransaminase$)).tw.
56. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
57. sgot.tw.
58. 55 or 56 or 57
59. 38 and 39 and 42 and 43 and 47 and 54 and 58
60. 38 or 39 or 42 or 43 or 47 or 54 or 58
61. 60 and 19
62. exp "Sensitivity and Specificity"/
63. sensitivity.tw.
64. specificity.tw.
65. ((pre-test or pretest) adj probability).tw.
66. post-test probability.tw.
67. predictive value$.tw.
68. likelihood ratio$.tw.
69. or/62-68
70. 27 and 69
71. 35 and 69
72. 50 and 69
73. 61 and 69
74. 70 or 71 or 72 or 73
75. iqur.tw.
76. biopredictive.tw.
77. echosens.tw.
78. 75 or 76 or 77
79. 8 or 16 or 33 or 48 or 59 or 74 or 78
80. "Quality of Life"/
81. (qol or (quality adj2 life)).ab,ti.
82. (value adj2 (money or monetary)).tw.
83. value of life/
84. quality adjusted life year/
85. quality adjusted life.tw.
86. (qaly$or qald$or qale$or qtime$).tw.
87. disability adjusted life.tw.
88. daly$.tw.
89. health status indicators/
90. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shorform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw.
91. (sf 6 or sf6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw.
92. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw.
93. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortfrom sixteen or short form sixteen).tw.
94. (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).tw.
95. (euroqol or euro qol or eq5d or eq 5d).tw.
96. (hql or hqol or h qol or hrqol or hr qol).tw.
97. (hye or hyes).tw.
98. health$year$equivalent$.tw.

99. health utilit$.tw.
100. (hui or hui1 or hui2 or hui3).tw.
101. disutilit$.tw.
102. rosser.tw.
103. (quality adj2 wellbeing).tw.
104. qwb.tw.
105. (willingness adj2 pay).tw.
106. standard gamble$.tw.
107. time trade off.tw.
108. time tradeoff.tw.
109. tto.tw.
110. letter.pt.
111. editorial.pt.
112. comment.pt.
113. 110 or 111 or 112
114. or/80-109
115. 114 not 113
116. exp liver cirrhosis/or exp liver diseases, alcoholic/
117. (fibros*s or cirrhos*s).tw.
118. 116 or 117
119. 115 and 118
120. (pulmonary or cystic).tw.
121. 119 not 120
122. 79 and 115
123. 122 or 121

# Appendix 6

# QUADAS: details of criteria for scoring studies

**1. Was the spectrum of patients representative of the patients who will receive the test in practice?**

| | |
|---|---|
| Yes | All patients had suspected ALD, and patients were recruited both prospectively and consecutively |
| No | Some patients were known to have liver fibrosis or cirrhosis, or patients were studied retrospectively or non-consecutively |
| Unclear | Insufficient details given about stage or recruitment methods to make a judgement about whether or not the patient spectrum would be scored 'yes' |

**2. Is the reference standard likely to correctly classify the target condition?**

| | |
|---|---|
| Yes | The reference standard was liver biopsy, HVPG measurement or endoscopy for oesophageal varices. The study excluded patients with liver biopsies < 10 mm in length[195] |
| No | Some or all patients received a different reference standard. The study included patients with liver biopsies < 10 mm in length[195] |
| Unclear | Reference standard is not stated; for liver biopsy, no data given on length of specimen or portal tracts[113] |

**3. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the performance of the two tests?**

| | |
|---|---|
| Yes | Reference standard was performed within 2 weeks of the index test |
| No | Reference standard was performed more than 2 weeks before or after the index test |
| Unclear | The time between reference standard and index test is not stated |

**4. Did the whole sample, or a random selection of the sample, receive verification using a reference standard?**

| | |
|---|---|
| Yes | All patients, or a random selection of patients, who received the index test went on to receive verification of their disease status using a reference standard |
| No | Some patients who received the index test did not receive verification of their true disease state and the selection of patients to receive the reference standard was not random |
| Unclear | This information is not reported |

**5. Did patients receive the same reference standard regardless of the index test result?**

| | |
|---|---|
| Yes | Selection of reference standard was not determined by the index text result |
| No | Selection of reference standard was determined by the index test result |
| Unclear | It is not clear whether or not selection of reference standard was determined by the index test result |

**6. Were the reference standard results interpreted without knowledge of the results of the index test?**

**7. Were the index test results interpreted without knowledge of the results of the reference standard?**

| | |
|---|---|
| Yes | The index test was interpreted without knowledge of the results of the reference standard or vice versa. If the test was clearly interpreted before the results of the other test were available then this was scored as 'yes' |
| No | The person interpreting the index test was aware of the results of the reference standard or vice versa |
| Unclear | No information is provided regarding whether or not tests were interpreted blindly |

**8. Were uninterpretable/intermediate test results reported and included in the analysis?**

| | |
|---|---|
| Yes | There were uninterpretable/intermediate results and they were included in the analysis, or patients were recruited prospectively and consecutively and no uninterpretable/intermediate results were reported |
| No | There were uninterpretable/intermediate results and they were excluded from the analysis |
| Unclear | It is not clear whether or not there were any uninterpretable/intermediate test results (this includes studies which were not prospective or not consecutive which did not report these data) |

| | |
|---|---|
| **9. Were withdrawals from the study explained?** | |
| Yes | All patients recruited to the study were accounted for |
| No | There appear to be patients who were recruited into the study who are not accounted for |
| Unclear | It is not clear whether or not any withdrawals occurred |
| **10. Were the selection criteria clearly described?** | |
| Yes | All relevant information about how patients were selected for inclusion in the study was provided |
| No | Study selection criteria were not clearly reported |
| Unclear | Insufficient details given about stage or recruitment methods to make a judgement about whether or not the selection criteria would be scored 'yes' |
| **11. Was the execution of the reference standard described in sufficient detail to permit replication of the test?** | |
| Yes | Sufficient details of the text execution are reported |
| No | Sufficient details are not reported |
| Unclear | Staging system given, but no inclusion criteria regarding the length of the liver biopsy or the number of portal tracts[113] |
| **12. Was the execution of the index test described in sufficient detail to permit replication of the test?** | |
| Yes | Sufficient details of the text execution are reported |
| No | Sufficient details are not reported |
| Unclear | Not applicable |
| **13. Did the study appear independent of the test manufacturer?** | |
| Yes | Study authors state that they have no personal interests in the company |
| No | One or more of study authors known to have an interest in the company |
| Unclear | No information provided |

# Appendix 7

# Excluded studies

Table of studies identified by the electronic searches and other searches, and excluded at the full paper stage for reasons not immediately apparent from the full text.

| Study | Reason for exclusion |
|---|---|
| Calès 2002[196] | Superseded by Calès 2005[197] |
| Calès 2005[197] | Clarification from the author that the validating population, in which FibroTest was used, was wholly composed of patients with hepatitis C (Dr Paul Calès, Université d'Angers, France, 2010, personal communication) |
| Clevert 2009[198] | Not available within study timescale |
| De Ledinghen 2006[199] | Information provided identical to that in Foucher 2006a[200] |
| Foucher 2006a[200] | Biopsy used only in a subset of 60 of patients in the study. The selection criteria for biopsy were not clear as 149 patients were said to have FibroScan results suggestive of cirrhosis |
| Foucher 2006b[201] | Clarification from the author that the population overlaps with that of Foucher 2006a[200] (Dr Juliette Foucher, Hôpital Haut-Lévèque, Pessac, France, 2010, personal communication), which reports data relating to substantially more patients with ALD; it is not clear how many patients with ALD underwent biopsy, and the selection criteria for biopsy were not clear, being described only as the 'usual indications for liver biopsy', and not in terms of the results of the FibroScan test |
| Jimenez-Ridruejo 2008[202] | Not available within study timescale |
| Laharie 2006[203] | Information provided identical to that in Foucher 2006a[200] |
| Lee 2009[204] | Not available within study timescale |
| Lieber 2008[62] | Although this study relates to the serum markers used in the ELF test, the tests do not appear to have been performed on an Immuno 1 machine, and therefore performance would be suboptimal (Professor William Rosenberg, University College London, 2010, personal communication) |
| Marin-Gabriel 2008[205] | Not available within study timescale |
| Melin 2005[206] | Superseded by Melin 2005[122] |
| Morozov 2008[207] | In Russian; no English abstract available |
| Mueller 2008[120] | Superseded by Mueller 2010[118] |
| Mueller 2009[208] | Superseded by Mueller 2010[118] |
| Nahon 2007[131] | Superseded by Nahon 2008[128] |
| Naveau 2003[209] | Appears to be superseded by Naveau 2005[123] (confirmation could not be obtained from the author) |
| Naveau 2008[210] | Superseded by Naveau 2009[124] |
| Nguyen-Khac 2007[211] | Superseded by Nguyen-Khac 2008[132] |
| Parkes 2007[212] | Superseded by Parkes 2010[57] |
| Rosenberg 2001[213] | Not available within study timescale |
| Thabut 2003[214] | Data relating specifically to patients with ALD neither published nor available from study authors (Dr Thierry Poynard, Hôpital Pitié-Salpêtrière, Paris, 2010, personal communication) |

# Appendix 8

# Diagnostic venepuncture: systematic review of adverse events

A systematic review was carried out in order to identify studies reporting adverse events in adults undergoing simple venepuncture for diagnostic or screening purposes. Literature searches were performed in February 2010. Details of the electronic databases that were searched, the search strategies used and the inclusion and exclusion criteria may be found in *Appendix 9*.

Studies that related to blood donors were excluded because:

- The withdrawal of larger volumes of blood makes it difficult to differentiate between vasovagal reactions and transient relative hypotension due to blood loss.[215]
- The use of needles with a larger bore than the 20–22 gauge generally used in blood sampling may increase the risk of injury,[216] as may the fact that the needles are in place for a longer period of time.[217]

Studies that used more invasive methods of blood collection (cannulation or catheterisation), or which collected arterial or capillary rather than venous blood samples, were also excluded.

The searches were not restricted by study type, language or date. The same three-stage sifting process was used as in the systematic review of clinical effectiveness. Data were extracted directly to the tables included in the report. Because many of the relevant studies took the form of case reports, a formal quality assessment was not undertaken. However, larger studies (observational or before-and-after studies) were deemed to be of higher quality than case series or case reports, and the latter were included only if they related to adverse events for which data were not available from the larger studies.

The electronic literature searches identified 979 potentially relevant articles. Of these, 11 met the review's inclusion criteria (for PRISMA diagram, see *Appendix 9*). These articles were:

- two observational studies, by Galena[215] and Deacon and Abramowitz[218]
- an uncontrolled before-and-after study by Godwin *et al.*[219]
- eight case reports by Choffel *et al.*,[220] Nouri *et al.*,[221] Pradhan and Gupta,[222] Saeed and Gatens,[223] Sander *et al.*,[224] Stitik *et al.*,[217] Vidal *et al.*[225] and Zubairy.[226]

A possibly relevant article by Rodriguez Guerrero *et al.*[227] was excluded because it was published in Spanish.

Five additional relevant articles, by Berry and Wallis,[228] Burgdorf *et al.*,[229] Horowitz,[216] Norcross and Shackford[230] and Yuan and Cohen,[231] were identified from citations.

The adverse events identified by the included studies fall into three major categories:

- vasovagal reactions
- pain and bruising
- nerve injuries.

These are discussed in turn below, as are the few studies relating to miscellaneous adverse events which do not fit into those categories.

## Vasovagal reactions

Vasovagal reactions result from an abnormal reflex stimulation of the vagus nerve. The trigger factors may be emotional or somatic.[232] In most patients, the signs and symptoms (which may include pallor, sweating, nausea, dizziness or light-headedness) are light or moderate, and resolve spontaneously. However, some patients experience bradycardia with consequent hypotension, loss of consciousness, and, in very severe cases, death.[232] In the context of vasovagal reactions associated with venepuncture, it seems likely that the greatest risk is that, when blood is taken with the patient sitting or lying, resumption of the upright position results in a faint, as the subsequent fall may result in injury.

Because data relating to vasovagal reactions were available from two large observational studies, lower-quality studies (case reports and small case series) relating to such adverse events were excluded.

The larger observational study, that by Galena,[215] recorded adverse effects associated with venepuncture carried out in outpatient settings between October 1988 and April 1991 on 4050 patients who were applying for life insurance. A 20- or 22-gauge needle was used to obtain a maximum of 30 ml blood from each patient. Delayed reactions were identified using telephone calls made an unspecified length of time after the venepuncture. Potentially serious vasovagal reactions were experienced by 3.4% of patients (*Table 39*); these were significantly more common in men than in women (4.0% vs 1.3%, $p < 0{,}001$). None of those who experienced convulsive syncope had a previous history of seizure disorder.

Deacon and Abramowitz[218] found lower rates of vasovagal reactions in 3315 adults undergoing venepuncture in three hospital outpatient phlebotomy clinics over a 3-week period, even though 80% of their population had fasted prior to venepuncture (*Table 40*). Although the rate of vasovagal reactions indicated by the phlebotomists was higher, at 0.9%, than that reported by the patients, it was still substantially lower than the rate of 3.4% reported by Galena.[215]

## Pain and bruising

Data relating to pain and bruising were available from one large observational study[215] and one uncontrolled before-and-after study.[219] Lower-quality studies (case reports and small case series) relating to such adverse events were therefore excluded.

**TABLE 39** Vasovagal reactions in patients undergoing venepuncture in outpatient settings[215]

| Complication | Number (%, 95% CI) |
|---|---|
| Diaphoresis, near syncope | 105/4050 (2.6, *2.1 to 3.1*) |
| Syncope | 24/4050 (0.6, *0.4 to 0.8*) |
| Convulsive syncope | 6/4050 (0.10, *0.03 to 0.30*) |
| Ventricular tachycardia | 1/4050 (*0.02, 0.00 to 0.10*) |
| Total | 136/4050 (3.4, *2.8 to 3.9*) |

Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

**TABLE 40** Vasovagal reactions in patients undergoing venepuncture in hospital phlebotomy clinics[218]

| Complication | Number (%, 95% CI) |
|---|---|
| Patient reported feeling very or extremely faint | 13/3315 (0.4, *0.2 to 0.6*) |
| Patient reported losing consciousness | 7/3315 (0.2, *0.1 to 0.4*) |
| Phlebotomist reported using strategies to manage fainting symptoms[a] with patient | *30/3315* (0.9, *0.6 to 1.2*) |

a For example, reclining the patient's chair, asking the patients to place their heads between their legs or using a cold towel.
Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

In Galena's large observational study,[215] 14.2% of patients reported adverse events related to pain and bruising (*Table 41*). Such adverse effects were significantly more common in women than in men (38.1% vs 7.9%, $p < 0{,}001$), a result that Galena suggested was probably related to their narrower veins. No cases of local cellulitis or phlebitis were reported.

Godwin *et al.*[219] reported higher overall rates of bruising in a small before-and-after study that audited bruising in two groups of 100 consecutive medical and surgical inpatients aged ≥ 15 years who were not receiving anticoagulants and did not have extensive pre-existing bruises. Venepuncture was performed by phlebotomists using a pre-evacuated tube collection system to take blood from the antecubital fossa. A clean cotton wool ball was then taped to the venepuncture site. The phlebotomist instructed patients in the first group to apply pressure for a few minutes after the venepuncture, but remained with patients in the second group until the bleeding had stopped. The venepuncture site was then assessed 24 hours later. Bruising was less common in the second group (45% vs 25%, $p < 0.01$) and such bruises as occurred were also smaller in this group. The difference between the groups was more marked in older patients (*Table 42*) and the investigators suggested that this was perhaps because they were less able than younger patients to apply pressure to the venepuncture site;[219] however, it is perhaps more likely to reflect the more fragile nature of the skin in the elderly.

## Nerve injury

The potentially most serious adverse events associated with venepuncture relate to nerve injury. Such adverse events can have disabling consequences. The only identified publications that report venepuncture-associated nerve injuries sufficiently severe to be brought to medical attention take the form of case reports and one small case series.[216]

The case series presented data relating to 11 patients who were referred to a specialist with a particular interest in nerve injuries because of causalgia following routine venepuncture.[216] However, only four of these patients had undergone venepuncture for blood sampling; in the remainder, the venepuncture was for blood donation, insertion of intravenous lines or intravenous medication. A later paper by Horowitz,[233] which combined data relating to these 11 patients with data from 13 patients who had subsequently been evaluated, could not be utilised because it presented aggregated data from patients who had undergone venepuncture for blood sampling and patients who had undergone venepuncture for other reasons.

Data relating to the cases identified in the case reports, together with the four relevant patients from Horowitz's[233] case series, are summarised in *Table 43*. These data demonstrate that nerve damage consequent on venepuncture can cause long-lasting pain, and loss of muscle power

**TABLE 41** Pain and bruising in patients undergoing venepuncture in outpatient settings[215]

| Complication | Number (%, 95% CI) |
|---|---|
| Bruising | 416/4050 (10.3, *9.3 to 11.2*) |
| Haematoma | 80/4050 (2.0, *1.6 to 2.4*) |
| Pain | 80/4050 (2.0, *1.6 to 2.4*) |
| Total | 576/4050 (14.2, *13.1 to 15.3*) |

Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

**TABLE 42** Bruising after venepuncture by haemostasis technique and patient age[219]

| | Number of patients with bruising (%, 95% CI) | |
|---|---|---|
| Patient age (years) | Patient pressure | Phlebotomist pressure |
| <60 | 11/37 (30, *15 to 44*) | 7/42 (17, *5 to 28*) |
| >60 | 34/63 (54, *42 to 66*) | 18/58 (31, *19 to 43*) |
| Total | 45/100 (45, *35 to 55*) | 25/100 (25, *17 to 33*) |

Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

and manual dexterity; it may also lead to clinical depression. Relatively few details of the venepuncture are reported, but in 4 of the 12 cases it was specifically said to have been difficult. The gauge of needle was specified in only two cases; in both cases, it was a 20-gauge needle.

The case studies summarised above do not provide any indication of the incidence of nerve injuries related to venepuncture, other than that they were rare. Some indication of the incidence can be obtained by considering only two studies from blood transfusion centres. In a New Zealand blood transfusion unit performing approximately 80,000 venepunctures a year, Berry and Wallis[228] found that, over a 2-year period, six people suffered injuries to the median nerve or medial and lateral cutaneous nerves that were severe enough for them to seek medical attention – an overall rate of approximately 1 in 25,000 (0.004%). Of those six, only one (summarised in *Table 43*) was undergoing venepuncture for diagnostic purposes, using a 20-gauge needle; the remaining five were undergoing venepuncture for blood donation, using a larger 16-gauge needle. As this study gave no indication of the number or proportion of venepunctures undertaken for purposes of diagnosis rather than blood donation, it is not possible to calculate a rate of nerve injury specific to diagnostic venepuncture; however, it seems likely that it would be lower than the overall rate.

Newman and Waxman[234] reported a higher nerve injury rate from a blood donation centre in the USA where nurses routinely reported all donor injuries. Over a 2-year period, 419,000 blood donations were collected using a 16-gauge needle and 66 cases of neurological nerve injury were identified from nursing records – a rate of 1 in 6300 (0.016%). This figure is not directly comparable with the New Zealand figure[228] because it includes cases that were not brought to medical attention, but the data for donors who requested a physician consultation (17 of the 56 individuals with nerve injury for whom follow-up data were available) also indicate a rate of approximately 1 in 25,000 (0.004%) (*Table 44*). This is a conservative estimate as 9 of the 66 donors with nerve injury could not be contacted for telephone follow-up, and one was deliberately not contacted because of pending litigation.[234]

DOI: 10.3310/hta16040



**TABLE 43** Nerve damage associated with venepuncture

| Study | Subject | Purpose and site of venepuncture | Diagnosis | Outcome | Comment |
|---|---|---|---|---|---|
| Berry 1977[228] | 50-year-old woman | Blood grouping; left antecubital fossa | Injury to the medial cutaneous nerve | Pain and swelling in the forearm developed within 24 hours into hyperaesthesia in the whole forearm. A striking improvement was noted 24 hours after treatment with carbamazepine and 3 days later the only symptom was slight pain on moving the arm. Treatment was discontinued after 5 weeks, when the patient had no symptoms except slightly impaired touch sensation in the sensory distribution of the left medial cutaneous nerve | A 20-gauge needle used |
| Horowitz 1994[216] | 61-year-old woman | Blood sampling; antecubital fossa | Causalgia affecting medial antebrachial cutaneous nerve | Increased symptoms and motor abnormalities of disuse, with joint contracture and psychiatric depression requiring antidepressant medication, observed at 7 years | |
| Horowitz 1994[216] | 61-year-old man | Blood sampling; antecubital fossa | Causalgia affecting lateral antebrachial cutaneous nerve | Increased symptoms and motor abnormalities of disuse, with joint contracture and psychiatric depression requiring antidepressant medication observed at 4 years | |
| Horowitz 1994[216] | 56-year-old woman | Blood sampling; antecubital fossa | Causalgia affecting medial antebrachial cutaneous nerve | Increased symptoms, with joint contracture and motor abnormalities of disuse observed at 18 months | |
| Horowitz 1994[216] | 35-year-old man | Blood sampling; wrist | Causalgia affecting superficial radial nerve | The burning pain resolved spontaneously over a 2-week period, but hyperpathia and allodynia in the injured nerve distribution persisted at 2.5 years | |
| Nouri 2000[221] | 59-year-old woman | Routine phlebotomy for pre-operative assessment; radial vein | Causalgia affecting radial nerve | Immediate acute pain and numbness; dysaesthesia, hyperaesthesia, allodynia, and loss of muscular power still persisted a year later. Following treatment with paroxetine, tramadol and capsaicin, and six nerve blocks, the pain in the arm and forearm was almost completely resolved, and that in the hand and wrist was somewhat reduced | A 20-gauge needle used. Venepuncture said to be difficult, requiring three attempts |
| Pradhan 1995[222] | 32-year-old woman with a minor pyrexial illness | Routine blood testing; cubital vein | Median nerve | Immediate intense pain in whole of left arm persisting on the palmar aspect of the forearm and hand, and accompanied by weakness and tingling. The paraesthesia subsided in 2 months; mild anaesthesia in radial side of palm persisted for 4 months; muscle power returned to normal with physiotherapy, but minimal wasting was still observed after 1 year | Venepuncture said to be very difficult because of the non-visibility of veins |

*continued*

**TABLE 43** Nerve damage associated with venepuncture *(continued)*

| Study | Subject | Purpose and site of venepuncture | Diagnosis | Outcome | Comment |
|---|---|---|---|---|---|
| Saeed 1983[223] | 47-year-old man | Preoperative phlebotomy; cubital vein | Anterior interosseous syndrome | Pain in forearm and inability to flex thumb noted 4 days after surgery. Surgical tendon transfer required 14 months later to enable appropriate movement of the thumb. | Venepuncture said to have been very difficult |
| Sander 1998[224] | 64-year-old woman | Phlebotomy (purpose not stated); antecubital | Lateral antebrachial cutaneous neuropathy | Acute pain on insertion of needle followed by pain and numbness persisting, with some improvement, for 5 months | |
| Stitik 2001[217] | 29-year-old man | Phlebotomy (purpose not stated); left cephalic vein, antecubital fossa | Lateral antebrachial cutaneous neuropathy | Shooting pain down forearm and into base of thumb at venepuncture; this initially resolved, but recurred the next morning together with dysaesthesias over the lateral aspect of the left distal forearm. At 6 months, dysaesthesias were reported in the entire left arm. The patient was then lost to follow-up | |
| Yuan 1985[231] | 31-year-old man | Routine phlebotomy for preoperative blood tests; cubital vein | Laceration of the lateral antebrachial cutaneous nerve with neuroma formation | Excruciating pain followed by numbness noted during venepuncture, followed by pain and numbness in the forearm persisting for 3 weeks, and resistant to treatment with butazolidin; lidocaine and steroid injection did not produce lasting relief. Surgery was performed on two occasions; the first was ineffective and the second relieved the pain, but left permanent numbness. However, motor function was unimpaired | Repeated attempts at venepuncture were required |
| Zubairy 2002[226] | 44-year-old woman | Routine postoperative blood sampling; cubital fossa | Severe anterior interosseus nerve lesion | Loss of function in the thumb and index finger; weakness of pronation. Management was conservative. The first sign of spontaneous recovery was observed at 20 months and normal function at 34 months after the injury | |

**TABLE 44** Number of blood donors with nerve injury following venepuncture[234]

| Recovery period | Number of donors with nerve injury and follow-up data ($n=56$) (% of total, 95% CI) | Number requesting physician consultation(s) (% of category, 95% CI) | Number with residual neurological defect[a] (% of category, 95% CI) |
|---|---|---|---|
| <3 days | 22 *(39, 27% to 52%)* | 0 (0) | 0 (0) |
| 3–29 days | 17 *(30, 18% to 42%)* | 5 (29, *8% to 51%*) | 0 (0) |
| 1–3 months | 13 *(4, 0% to 8%)* | 8 (62, *35% to 88%*) | 2 (*15, 0% to 35%*) |
| 3–6 months | 2 *(23, 12% to 34%)* | 2 (100) | 1 (50, *0% to 100%*) |
| >6 months | 2 *(23, 12% to 34%)* | 2 (100) | 1 (50, *0% to 100%*) |

a Mild localised numbness which did not interfere with function.
Data in roman were taken directly from the text; data in *italics* were calculated by the reviewer.

## Miscellaneous adverse events

Three case reports highlight the ability of venepuncture to provoke localised manifestations of underlying medical conditions. Burgdorf *et al.*[229] reported multiple sarcoid granulomas that developed following numerous venepunctures for diagnostic purposes in a 38-year-old woman who was subsequently diagnosed with sarcoidosis. Similarly, Choffel *et al.*[220] reported the formation of a sarcoid granuloma at the puncture site following venepuncture in a 56-year-old woman with sarcoidosis. Finally, Vidal *et al.*[225] reported that a tuberculous ulcerated nodule developed on the left wrist of an 88-year-old man with a history of pulmonary tuberculosis 2 years after venepuncture at that site; they considered the venepuncture to be the most likely cause of the reactivation of tuberculous infection.

Finally, a 27-year-old woman was aware of a 'buzzing' in the region of her left antecubital fossa following venepuncture performed by a gynaecologist rather than a medical technician. This buzzing became more pronounced over a 2-year period, and was eventually identified during a routine medical examination as an arteriovenous fistula which required surgical repair.[230]

## Summary

There is evidence that venepuncture may be associated with adverse effects. Vasovagal reactions were studied in two large observational studies, by Galena[215] and Deacon and Abramowitz,[218] which together included 7365 individuals. Data relating to pain and bruising were available from a total of 4250 patients included in Galena's[215] large observational study and the small before-and-after study by Godwin *et al.*[219] Unfortunately, data relating to direct nerve injuries in patients undergoing venepuncture specifically for diagnostic or screening purposes were available from only case series or case reports.

The most commonly reported adverse events were those related to pain and bruising, which affected between 14% and 45% of patients. Vasovagal reactions were rarer, affecting between 0.9% and 3.4%. There were no data regarding the incidence of nerve injuries associated with diagnostic venepuncture, but it seems likely that it would be lower than the 0.004% reported in blood donors. However, although such nerve injuries appear to be very rare, they are potentially disabling.

# Appendix 9

# Systematic review of the adverse effects of venepuncture: search strategies

## Sources searched

The electronic bibliographic databases which were searched are listed in *Appendix 5*.

## Search strategies

The MEDLINE search strategy may be found in *Appendix 5*.

### *Inclusion criteria*

#### Population

- Adults.

#### Intervention

- Simple venepuncture for diagnostic or screening purposes.

#### Outcomes

- Adverse events probably or possibly caused by the process of testing.

#### Setting

- Any country.

#### Study type

- RCTs.
- Controlled non-randomised studies (egg cohort studies).
- Case–control studies.
- Case series.
- Case reports.
- Systematic reviews.
- Economic evaluations.

### *Exclusion criteria*

#### Population

- Studies relating specifically to people receiving anticoagulation therapy, as their propensity to bruise would be significantly greater than that of patients not receiving anticoagulation therapy.

## Intervention

Studies in which:

- venepuncture was used either specifically to obtain blood donations, or to obtain both blood donations and smaller samples for diagnostic or screening purposes, but did not present separate data relating to the two uses
- cannulation or catheterisation was used to obtain blood samples
- the study related to the collection of arterial or capillary rather than venous blood samples.

## Study type

- Animal models.
- Narrative reviews, editorials, opinions.

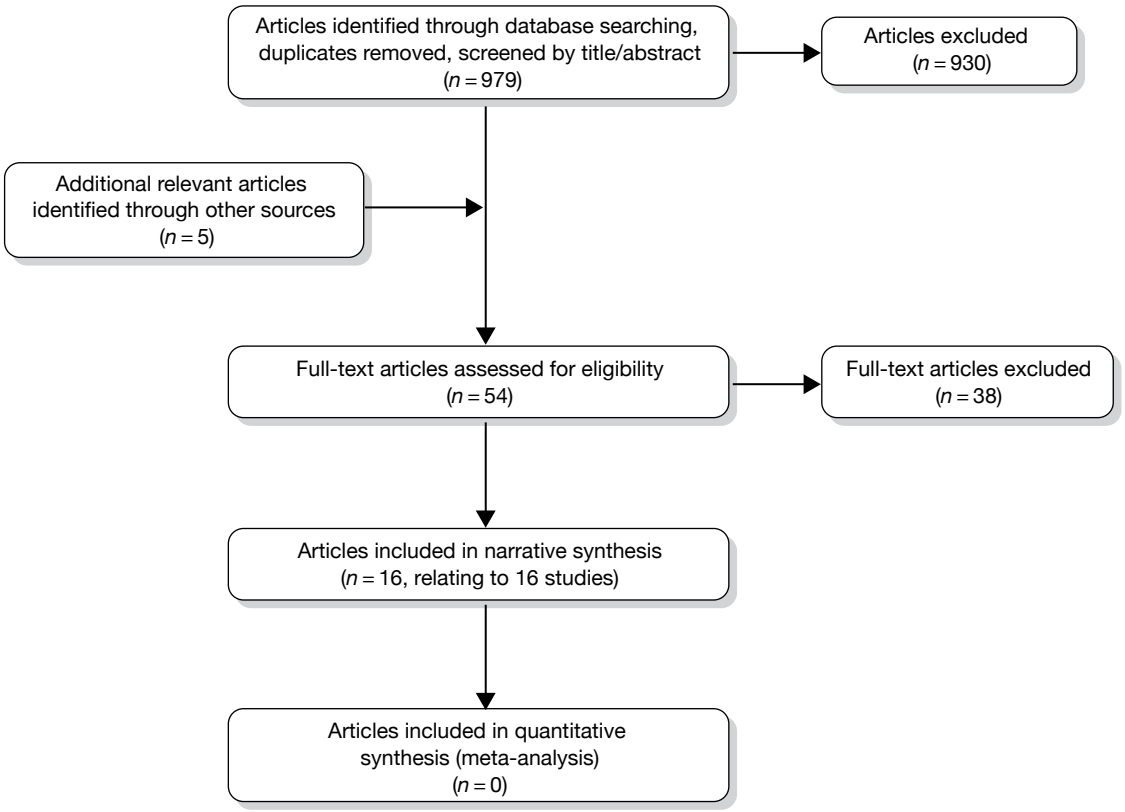



**FIGURE 29** Adverse effects of venepuncture: summary of study selection and exclusion.

# Appendix 10

## The results from the cost-effectiveness analyses

*Table 45* provides the costs and QALYs from each strategy assuming that there is no benefit associated with biopsy, nor any change in abstinence rates associated with diagnostic test. *Table 46* provides the proportion of patients undergoing biopsy categorised by scenario number and diagnostic test.

*Table 47* provides the threshold level for percentage point decrease in abstinence rates that would make biopsy all the more cost-effective strategy compared with each NILT; *Table 48* provides the same detail, but the threshold value is the QALY gain associated with a biopsy

**TABLE 45** The costs and QALYs from the 36 scenarios

| | Biopsying all | | FibroScan triage | | ELF triage | | FibroTest triage | | Clinical experience triage | | FibroScan replacement | | ELF replacement | | FibroTest replacement | | Clinical experience replacement | | Diagnose all as cirrhotic | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scenario number | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs |
| 1 | 16,246 | 9.36 | 15,922 | 9.35 | 15,815 | 9.33 | 15,870 | 9.32 | 18,020 | 9.39 | 15,855 | 9.34 | 16,915 | 9.35 | 16,246 | 9.36 | 15,922 | 9.35 | 15,815 | 9.33 |
| 2 | 16,246 | 9.32 | 15,922 | 9.33 | 15,815 | 9.31 | 15,870 | 9.30 | 18,020 | 9.38 | 15,855 | 9.33 | 16,915 | 9.34 | 16,246 | 9.32 | 15,922 | 9.33 | 15,815 | 9.31 |
| 3 | 16,841 | 9.35 | 16,270 | 9.35 | 16,111 | 9.33 | 16,195 | 9.32 | 18,139 | 9.39 | 16,004 | 9.33 | 17,063 | 9.34 | 16,841 | 9.35 | 16,270 | 9.35 | 16,111 | 9.33 |
| 4 | 16,841 | 9.31 | 16,270 | 9.32 | 16,111 | 9.31 | 16,195 | 9.30 | 18,139 | 9.38 | 16,004 | 9.32 | 17,063 | 9.33 | 16,841 | 9.31 | 16,270 | 9.32 | 16,111 | 9.31 |
| 5 | 16,246 | 9.36 | 15,794 | 9.26 | 16,204 | 9.35 | 16,261 | 9.36 | 18,084 | 9.30 | 23,855 | 9.47 | 20,642 | 9.43 | 16,246 | 9.36 | 15,794 | 9.26 | 16,204 | 9.35 |
| 6 | 16,246 | 9.32 | 15,794 | 9.24 | 16,204 | 9.32 | 16,261 | 9.33 | 18,084 | 9.29 | 23,855 | 9.46 | 20,642 | 9.42 | 16,246 | 9.32 | 15,794 | 9.24 | 16,204 | 9.32 |
| 7 | 16,841 | 9.35 | 16,115 | 9.26 | 16,743 | 9.35 | 16,706 | 9.35 | 18,203 | 9.30 | 24,004 | 9.47 | 20,791 | 9.43 | 16,841 | 9.35 | 16,115 | 9.26 | 16,743 | 9.35 |
| 8 | 16,841 | 9.31 | 16,115 | 9.24 | 16,743 | 9.31 | 16,706 | 9.32 | 18,203 | 9.29 | 24,004 | 9.46 | 20,791 | 9.42 | 16,841 | 9.31 | 16,115 | 9.24 | 16,743 | 9.31 |
| 9 | 16,246 | 9.36 | 15,792 | 9.29 | 15,871 | 9.35 | 15,863 | 9.27 | 17,277 | 9.31 | 16,617 | 9.36 | 17,698 | 9.30 | 16,246 | 9.36 | 15,792 | 9.29 | 15,871 | 9.35 |
| 10 | 16,246 | 9.32 | 15,792 | 9.27 | 15,871 | 9.32 | 15,863 | 9.25 | 17,277 | 9.31 | 16,617 | 9.35 | 17,698 | 9.29 | 16,246 | 9.32 | 15,792 | 9.27 | 15,871 | 9.32 |
| 11 | 16,841 | 9.35 | 16,097 | 9.28 | 16,194 | 9.34 | 16,193 | 9.26 | 17,396 | 9.31 | 16,765 | 9.36 | 17,847 | 9.30 | 16,841 | 9.35 | 16,097 | 9.28 | 16,194 | 9.34 |
| 12 | 16,841 | 9.31 | 16,097 | 9.26 | 16,194 | 9.32 | 16,193 | 9.24 | 17,396 | 9.30 | 16,765 | 9.35 | 17,847 | 9.29 | 16,841 | 9.31 | 16,097 | 9.26 | 16,194 | 9.32 |
| 13 | 16,135 | 9.24 | 15,721 | 9.18 | 15,628 | 9.16 | 15,274 | 9.16 | 18,138 | 9.29 | 16,118 | 9.24 | 16,972 | 9.24 | 16,135 | 9.24 | 15,721 | 9.18 | 15,628 | 9.16 |
| 14 | 16,135 | 9.20 | 15,721 | 9.16 | 15,628 | 9.14 | 15,274 | 9.14 | 18,138 | 9.29 | 16,118 | 9.23 | 16,972 | 9.23 | 16,135 | 9.20 | 15,721 | 9.16 | 15,628 | 9.14 |
| 15 | 16,730 | 9.24 | 16,049 | 9.18 | 15,909 | 9.16 | 15,577 | 9.15 | 18,257 | 9.29 | 16,267 | 9.24 | 17,121 | 9.24 | 16,730 | 9.24 | 16,049 | 9.18 | 15,909 | 9.16 |
| 16 | 16,730 | 9.20 | 16,049 | 9.15 | 15,909 | 9.14 | 15,577 | 9.13 | 18,257 | 9.28 | 16,267 | 9.23 | 17,121 | 9.23 | 16,730 | 9.20 | 16,049 | 9.15 | 15,909 | 9.14 |
| 17 | 16,135 | 9.24 | 15,638 | 9.12 | 16,127 | 9.22 | 16,037 | 9.18 | 18,546 | 9.22 | 25,346 | 9.44 | 21,160 | 9.33 | 16,135 | 9.24 | 15,638 | 9.12 | 16,127 | 9.22 |
| 18 | 16,135 | 9.20 | 15,638 | 9.10 | 16,127 | 9.18 | 16,037 | 9.15 | 18,546 | 9.21 | 25,346 | 9.43 | 21,160 | 9.32 | 16,135 | 9.20 | 15,638 | 9.10 | 16,127 | 9.18 |

| | Biopsying all | | FibroScan triage | | ELF triage | | FibroTest triage | | Clinical experience triage | | FibroScan replacement | | ELF replacement | | FibroTest replacement | | Clinical experience replacement | | Diagnose all as cirrhotic | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scenario number | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs | Costs (£) | QALYs |
| 19 | 16,730 | 9.24 | 15,952 | 9.11 | 16,702 | 9.21 | 16,472 | 9.17 | 18,665 | 9.22 | 25,495 | 9.44 | 21,308 | 9.32 | 16,730 | 9.24 | 15,952 | 9.11 | 16,702 | 9.21 |
| 20 | 16,730 | 9.20 | 15,952 | 9.09 | 16,702 | 9.17 | 16,472 | 9.14 | 18,665 | 9.21 | 25,495 | 9.43 | 21,308 | 9.32 | 16,730 | 9.20 | 15,952 | 9.09 | 16,702 | 9.17 |
| 21 | 16,135 | 9.24 | 15,627 | 9.13 | 15,675 | 9.17 | 16,123 | 9.17 | 17,933 | 9.22 | 16,817 | 9.26 | 23,080 | 9.35 | 16,135 | 9.24 | 15,627 | 9.13 | 15,675 | 9.17 |
| 22 | 16,135 | 9.20 | 15,627 | 9.11 | 15,675 | 9.15 | 16,123 | 9.14 | 17,933 | 9.22 | 16,817 | 9.25 | 23,080 | 9.34 | 16,135 | 9.20 | 15,627 | 9.11 | 15,675 | 9.15 |
| 23 | 16,730 | 9.24 | 15,927 | 9.12 | 15,980 | 9.17 | 16,611 | 9.16 | 18,052 | 9.22 | 16,965 | 9.26 | 23,229 | 9.34 | 16,730 | 9.24 | 15,927 | 9.12 | 15,980 | 9.17 |
| 24 | 16,730 | 9.20 | 15,927 | 9.10 | 15,980 | 9.15 | 16,611 | 9.13 | 18,052 | 9.22 | 16,965 | 9.25 | 23,229 | 9.33 | 16,730 | 9.20 | 15,927 | 9.10 | 15,980 | 9.15 |
| 25 | 16,135 | 9.24 | 15,783 | 9.24 | 15,710 | 9.22 | 16,236 | 9.22 | 17,267 | 9.36 | 15,827 | 9.31 | 15,834 | 9.31 | 16,135 | 9.24 | 15,783 | 9.24 | 15,710 | 9.22 |
| 26 | 16,135 | 9.20 | 15,783 | 9.22 | 15,710 | 9.20 | 16,236 | 9.20 | 17,267 | 9.35 | 15,827 | 9.30 | 15,834 | 9.30 | 16,135 | 9.20 | 15,783 | 9.22 | 15,710 | 9.20 |
| 27 | 16,730 | 9.24 | 16,111 | 9.24 | 16,006 | 9.22 | 16,532 | 9.22 | 17,386 | 9.36 | 15,976 | 9.31 | 15,983 | 9.31 | 16,730 | 9.24 | 16,111 | 9.24 | 16,006 | 9.22 |
| 28 | 16,730 | 9.20 | 16,111 | 9.22 | 16,006 | 9.20 | 16,532 | 9.20 | 17,386 | 9.35 | 15,976 | 9.30 | 15,983 | 9.30 | 16,730 | 9.20 | 16,111 | 9.22 | 16,006 | 9.20 |
| 29 | 16,135 | 9.24 | 15,738 | 9.21 | 15,983 | 9.24 | 16,962 | 9.24 | 17,380 | 9.32 | 21,265 | 9.41 | 19,979 | 9.39 | 16,135 | 9.24 | 15,738 | 9.21 | 15,983 | 9.24 |
| 30 | 16,135 | 9.20 | 15,738 | 9.19 | 15,983 | 9.21 | 16,962 | 9.22 | 17,380 | 9.31 | 21,265 | 9.40 | 19,979 | 9.38 | 16,135 | 9.20 | 15,738 | 9.19 | 15,983 | 9.21 |
| 31 | 16,730 | 9.24 | 16,057 | 9.21 | 16,448 | 9.24 | 17,390 | 9.24 | 17,499 | 9.32 | 21,413 | 9.41 | 20,128 | 9.39 | 16,730 | 9.24 | 16,057 | 9.21 | 16,448 | 9.24 |
| 32 | 16,730 | 9.20 | 16,057 | 9.19 | 16,448 | 9.21 | 17,390 | 9.21 | 17,499 | 9.31 | 21,413 | 9.40 | 20,128 | 9.38 | 16,730 | 9.20 | 16,057 | 9.19 | 16,448 | 9.21 |
| 33 | 16,135 | 9.24 | 15,635 | 9.20 | 15,749 | 9.23 | 15,526 | 9.18 | 15,595 | 9.28 | 16,383 | 9.33 | 16,350 | 9.26 | 16,135 | 9.24 | 15,635 | 9.20 | 15,749 | 9.23 |
| 34 | 16,135 | 9.20 | 15,635 | 9.18 | 15,749 | 9.21 | 15,526 | 9.16 | 15,595 | 9.27 | 16,383 | 9.32 | 16,350 | 9.25 | 16,135 | 9.20 | 15,635 | 9.18 | 15,749 | 9.21 |
| 35 | 16,730 | 9.24 | 15,895 | 9.19 | 16,065 | 9.23 | 15,818 | 9.17 | 15,714 | 9.28 | 16,532 | 9.32 | 16,499 | 9.25 | 16,730 | 9.24 | 15,895 | 9.19 | 16,065 | 9.23 |
| 36 | 16,730 | 9.20 | 15,895 | 9.18 | 16,065 | 9.21 | 15,818 | 9.15 | 15,714 | 9.27 | 16,532 | 9.31 | 16,499 | 9.24 | 16,730 | 9.20 | 15,895 | 9.18 | 16,065 | 9.21 |

**TABLE 46** The proportion of patients undergoing biopsy

| Scenario numbers | Biopsying all (%) | FibroScan triage (%) | ELF triage (%) | FibroTest triage (%) | Clinical experience triage (%) |
|---|---|---|---|---|---|
| 1–4 | 100 | 49 | 34 | 35 | 40 |
| 5–8 | 100 | 43 | 88 | 35 | 67 |
| 9–12 | 100 | 40 | 40 | 35 | 41 |
| 13–16 | 100 | 45 | 30 | 35 | 35 |
| 17–20 | 100 | 41 | 96 | 35 | 64 |
| 21–24 | 100 | 39 | 36 | 35 | 76 |
| 25–28 | 100 | 45 | 34 | 35 | 34 |
| 29–32 | 100 | 43 | 71 | 35 | 63 |
| 33–36 | 100 | 31 | 38 | 35 | 33 |

Note that for all scenarios evaluated for the replacement strategies no patients would receive a biopsy.

**TABLE 47** The threshold level (in percentage points) for the decrease in abstinence rate

| Scenario number | FibroScan triage | ELF triage | FibroTest triage | Clinical experience triage | FibroScan replacement | ELF replacement | FibroTest replacement | Clinical experience replacement | Diagnose all as cirrhotic |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.017 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.039 | 0.015 | 0.005 | 0.000 | 0.000 | 0.012 | 0.000 | 0.000 | 0.000 |
| 3 | 0.039 | 0.015 | 0.005 | 0.000 | 0.000 | 0.012 | 0.000 | 0.000 | 0.000 |
| 4 | 0.062 | 0.036 | 0.025 | 0.000 | 0.003 | 0.024 | 0.006 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 6 | 0.000 | 0.013 | 0.020 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 7 | 0.000 | 0.013 | 0.021 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 8 | 0.000 | 0.034 | 0.043 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 9 | 0.000 | 0.008 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 10 | 0.000 | 0.030 | 0.000 | 0.000 | 0.000 | 0.007 | 0.000 | 0.000 | 0.000 |
| 11 | 0.000 | 0.030 | 0.000 | 0.000 | 0.000 | 0.006 | 0.000 | 0.000 | 0.000 |
| 12 | 0.000 | 0.052 | 0.000 | 0.000 | 0.000 | 0.019 | 0.000 | 0.000 | 0.000 |
| 13 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 | 0.000 |
| 14 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.011 | 0.000 | 0.016 | 0.000 |

| Scenario number | FibroScan triage | ELF triage | FibroTest triage | Clinical experience triage | FibroScan replacement | ELF replacement | FibroTest replacement | Clinical experience replacement | Diagnose all as cirrhotic |
|---|---|---|---|---|---|---|---|---|---|
| 15 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.011 | 0.000 | 0.015 | 0.000 |
| 16 | 0.000 | 0.000 | 0.000 | 0.005 | 0.005 | 0.023 | 0.006 | 0.028 | 0.000 |
| 17 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 | 0.000 |
| 18 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.016 | 0.000 |
| 19 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.015 | 0.000 |
| 20 | 0.000 | 0.000 | 0.000 | 0.005 | 0.000 | 0.000 | 0.000 | 0.028 | 0.000 |
| 21 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 | 0.000 |
| 22 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.005 | 0.000 | 0.016 | 0.000 |
| 23 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.005 | 0.000 | 0.015 | 0.000 |
| 24 | 0.000 | 0.000 | 0.000 | 0.005 | 0.000 | 0.018 | 0.000 | 0.028 | 0.000 |
| 25 | 0.022 | 0.001 | 0.000 | 0.000 | 0.025 | 0.034 | 0.035 | 0.003 | 0.000 |
| 26 | 0.045 | 0.022 | 0.000 | 0.000 | 0.039 | 0.047 | 0.047 | 0.016 | 0.000 |
| 27 | 0.045 | 0.022 | 0.000 | 0.000 | 0.039 | 0.046 | 0.047 | 0.015 | 0.000 |
| 28 | 0.067 | 0.043 | 0.017 | 0.005 | 0.052 | 0.059 | 0.060 | 0.028 | 0.000 |
| 29 | 0.000 | 0.018 | 0.000 | 0.000 | 0.008 | 0.000 | 0.000 | 0.003 | 0.000 |
| 30 | 0.007 | 0.041 | 0.000 | 0.000 | 0.022 | 0.000 | 0.000 | 0.016 | 0.000 |
| 31 | 0.007 | 0.041 | 0.000 | 0.000 | 0.022 | 0.000 | 0.000 | 0.015 | 0.000 |
| 32 | 0.027 | 0.063 | 0.000 | 0.005 | 0.036 | 0.000 | 0.007 | 0.028 | 0.000 |
| 33 | 0.000 | 0.009 | 0.000 | 0.000 | 0.027 | 0.030 | 0.001 | 0.003 | 0.000 |
| 34 | 0.001 | 0.030 | 0.000 | 0.000 | 0.042 | 0.043 | 0.014 | 0.016 | 0.000 |
| 35 | 0.001 | 0.030 | 0.000 | 0.000 | 0.041 | 0.042 | 0.014 | 0.015 | 0.000 |
| 36 | 0.020 | 0.052 | 0.003 | 0.005 | 0.055 | 0.055 | 0.026 | 0.028 | 0.000 |

**TABLE 48** The threshold level (in QALYs gained) for the benefits associated with a biopsy

| Scenario number | FibroScan triage | ELF triage | FibroTest triage | Clinical experience triage | FibroScan replacement | ELF replacement | FibroTest replacement | Clinical experience replacement | Diagnose all as cirrhotic |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.029 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.069 | 0.028 | 0.008 | 0.000 | 0.000 | 0.037 | 0.000 | 0.000 | 0.000 |
| 3 | 0.069 | 0.029 | 0.009 | 0.000 | 0.000 | 0.036 | 0.000 | 0.000 | 0.000 |
| 4 | 0.109 | 0.068 | 0.048 | 0.000 | 0.008 | 0.076 | 0.017 | 0.000 | 0.000 |
| 5 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 6 | 0.000 | 0.023 | 0.035 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 7 | 0.000 | 0.023 | 0.036 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 8 | 0.000 | 0.063 | 0.076 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 9 | 0.000 | 0.014 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 10 | 0.000 | 0.054 | 0.000 | 0.000 | 0.000 | 0.020 | 0.000 | 0.000 | 0.000 |
| 11 | 0.000 | 0.055 | 0.000 | 0.000 | 0.000 | 0.019 | 0.000 | 0.000 | 0.000 |
| 12 | 0.000 | 0.094 | 0.000 | 0.000 | 0.000 | 0.059 | 0.000 | 0.000 | 0.000 |
| 13 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.008 | 0.000 |
| 14 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.033 | 0.000 | 0.048 | 0.000 |
| 15 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.032 | 0.000 | 0.047 | 0.000 |
| 16 | 0.000 | 0.000 | 0.000 | 0.009 | 0.015 | 0.072 | 0.018 | 0.086 | 0.000 |
| 17 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.008 | 0.000 |
| 18 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.048 | 0.000 |

| Scenario number | FibroScan triage | ELF triage | FibroTest triage | Clinical experience triage | FibroScan replacement | ELF replacement | FibroTest replacement | Clinical experience replacement | Diagnose all as cirrhotic |
|---|---|---|---|---|---|---|---|---|---|
| 19 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.047 | 0.000 |
| 20 | 0.000 | 0.000 | 0.000 | 0.009 | 0.000 | 0.000 | 0.000 | 0.086 | 0.000 |
| 21 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.008 | 0.000 |
| 22 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.016 | 0.000 | 0.048 | 0.000 |
| 23 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.015 | 0.000 | 0.047 | 0.000 |
| 24 | 0.000 | 0.000 | 0.000 | 0.009 | 0.000 | 0.055 | 0.000 | 0.086 | 0.000 |
| 25 | 0.038 | 0.001 | 0.000 | 0.000 | 0.073 | 0.105 | 0.108 | 0.008 | 0.000 |
| 26 | 0.078 | 0.041 | 0.000 | 0.000 | 0.113 | 0.145 | 0.148 | 0.048 | 0.000 |
| 27 | 0.078 | 0.041 | 0.000 | 0.000 | 0.112 | 0.144 | 0.146 | 0.047 | 0.000 |
| 28 | 0.118 | 0.081 | 0.031 | 0.009 | 0.151 | 0.184 | 0.186 | 0.086 | 0.000 |
| 29 | 0.000 | 0.032 | 0.000 | 0.000 | 0.020 | 0.000 | 0.000 | 0.008 | 0.000 |
| 30 | 0.014 | 0.071 | 0.000 | 0.000 | 0.060 | 0.000 | 0.000 | 0.048 | 0.000 |
| 31 | 0.014 | 0.072 | 0.000 | 0.000 | 0.059 | 0.000 | 0.000 | 0.047 | 0.000 |
| 32 | 0.054 | 0.112 | 0.000 | 0.009 | 0.098 | 0.000 | 0.021 | 0.086 | 0.000 |
| 33 | 0.000 | 0.015 | 0.000 | 0.000 | 0.077 | 0.093 | 0.003 | 0.008 | 0.000 |
| 34 | 0.001 | 0.055 | 0.000 | 0.000 | 0.117 | 0.133 | 0.043 | 0.048 | 0.000 |
| 35 | 0.001 | 0.056 | 0.000 | 0.000 | 0.115 | 0.132 | 0.041 | 0.047 | 0.000 |
| 36 | 0.041 | 0.095 | 0.006 | 0.009 | 0.155 | 0.171 | 0.081 | 0.086 | 0.000 |

# Appendix 11

## Project protocol

**HTA Reference No. 09/62/01**

*Final amended version, 19 January 2010*

## 1. Title of the project

Non-invasive diagnostic assessment tools for the detection of liver fibrosis in patients with suspected alcohol-related liver disease

## 2. Name of Assessment Team and project lead

*Assessment Team*
ScHARR Technology Assessment Group, University of Sheffield.

*Project Lead*
Matt Stevenson, Senior Research Fellow, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA
Tel: 0114 222 0691, Fax: 0114 272 4095, E-mail: m.d.stevenson@sheffield.ac.uk

*Address for correspondence*
Major documentation should be sent to the project lead (m.d.stevenson@sheffield.ac.uk), the project administrator (Andrea Shippam, a.shippam@sheffield.ac.uk) and the managing director of ScHARR-TAG (Eva Kaltenthaler, e.kaltenthaler@sheffield.ac.uk).

## 3. Plain English Summary

Excess alcohol consumption is associated with alcoholic liver disease (ALD): alcoholic fatty liver (steatosis), alcoholic hepatitis, or alcoholic cirrhosis.[1] Steatosis, which usually asymptomatic, is reversible if alcohol consumption is stopped or significantly reduced.[1] Alcoholic hepatitis involves more severe liver damage.[2] Some patients are asymptomatic, but many suffer abdominal symptoms, and others present with acute alcoholic hepatitis characterised by jaundice, fever, liver failure, or bleeding.[1] In alcoholic cirrhosis, scar tissue (fibrosis) prevents the liver from working properly;[1] despite this, some people with early-stage alcoholic cirrhosis have no symptoms.[2] People with alcoholic cirrhosis are at increased risk of liver cancer.[1]

People who drink more than 10 units of alcohol daily will eventually develop steatosis; 10%-35% will develop alcoholic hepatitis, and approximately 10% will develop cirrhosis.[3] Some develop both cirrhosis and alcoholic hepatitis;[4] over 60% of these patients die within four years of diagnosis.[2] Abstinence from alcohol greatly improves survival in people with ALD.[1]

Patients with ALD come to medical attention in a number of ways. Many are identified following routine liver function tests, others when they report relatively mild abdominal symptoms. Some present with more severe symptoms caused by advanced liver disease.[3] Yet others present voluntarily for detoxification, require treatment for alcohol-related injuries, or present with alcoholic damage to other organs.[3] Liver biopsy may be used to confirm the diagnosis of ALD and provide information about the degree of fibrosis.[3,5] As an invasive procedure, it carries a risk of morbidity and mortality, particularly in patients with alcoholic hepatitis and cirrhosis.[3] Moreover, there is no high-quality evidence for its accuracy,[5] and therefore current draft guidance recommends that it is used only when confirmation of a diagnosis of acute alcoholic hepatitis is needed to inform specific treatment decisions.[5]

The key element of treatment for patients with ALD is long-term abstinence from alcohol. Other elements aim to prevent disease progression and manage complications. These include lifestyle changes (reducing smoking and obesity), nutritional therapy,[2] and therapies to treat specific complications of ALD.[3] Liver transplantation may be offered in extreme cases.[3]

At least 7,000 new cases of cirrhosis are diagnosed in the UK each year,[6] and in 2007 4,580 people in England and Wales died from ALD.[7] Around 80% of all cases of liver cirrhosis seen in district general hospitals in the UK are due to alcohol,[3] and many people in England and Wales consume alcohol at levels which put them at risk of ALD. In 2007, 24.2% of adults in England reported hazardous or harmful patterns of alcohol consumption.[8] Directly comparable figures are not available for Wales.[9]

The aim of this review is to systematically evaluate and appraise the potential clinical and cost effectiveness of using non-invasive liver assessment tools in patients who might otherwise be candidates for biopsy or referral to specialist care.

## 4. Decision problem

### *4.1 Purpose of the decision to be made*

The aim of the assessment is to answer the following research question: Will using non-invasive liver assessment tools in patients with suspected alcohol-related liver fibrosis who might otherwise be candidates for biopsy or referral to specialist care reduce the number of referrals or biopsies and improve the health outcomes and quality of life of those patients?

### *4.2 Clear definition of the intervention*

Four interventions are considered in this assessment: three are composite blood tests, and the fourth is a specialised scan.

The Enhanced Liver Fibrosis (ELF) test (iQur Ltd) is a blood test which uses an algorithm combining three biomarkers (hyaluronic acid, procollagen III amino terminal peptide and tissue inhibitor of metalloproteinase) to assess the stage and rate of progression of liver fibrosis. The

biomarkers are direct markers of extracellular matrix metabolism/degradation indicative of liver fibrosis. A higher concentration of the individual biomarkers leads to a higher ELF score, and therefore it is more likely there is more severe fibrosis. It is proposed that the ELF test can be used for the baseline determination of liver fibrosis. The ELF test is CE marked.

FibroTest and FibroMax (BioPredictive) are both proprietary algorithms of markers based on blood tests to assess the stage of liver fibrosis. FibroTest uses alpha-2 macroglobulin, a direct marker of extracellular matrix metabolism/degradation, and four indirect markers (apolipoprotein A1, haptoglobin, bilirubin, and gamma-glutamyl-transpeptidase). FibroMax adds to FibroTest additional markers for steatosis and alcohol related disease: these additional markers include ALT, AST, glucose, height and weight. Neither FibroTest nor FibroMax are CE marked, but there are CE marked kits for assessing the appropriate components.

FibroScan (EchoSens) is a device which uses transient elastography to assess liver stiffness, which is correlated with the degree of fibrosis. It consists of a specialised probe, an ultrasound and elastography system, and specialised software. The probe is placed on the skin over the liver, and generates a mechanical pulse which sends a shear wave through the liver. Liver stiffness is calculated from the velocity of the wave, which is measured by ultrasound. FibroScan is CE marked.

*4.3 Place of the intervention in the treatment pathway(s)*

The assessment will investigate the effect of using any of the four interventions in patients with suspected alcohol-related liver fibrosis who might otherwise be referred for biopsy or specialist care on the basis of their clinical history and physical examination and/or standard liver function tests. If data and resources allow, the effectiveness of tests in combination will also be assessed.

*4.4 Relevant comparators*

Referral to specialty care or biopsy based on clinical suspicion of liver fibrosis based on symptoms and/or liver function test results.

*4.5 Populations and relevant subgroups*

Patients with suspected liver fibrosis related to alcohol consumption. If time permits, consideration will be given to the subgroup of patients with suspected liver fibrosis who have hepatitis C in addition to high alcohol consumption.

*4.6 Key factors to be addressed*

The review will aim to:

- Investigate by systematic review the diagnostic accuracy of each of the four interventions in patients with suspected alcohol-related liver fibrosis

- Investigate by systematic review the impact of the four interventions on health and quality of life outcomes in patients with suspected alcohol-related liver fibrosis
- Estimate the potential benefits and harms arising from altered treatment based on the results of the four interventions
- Estimate the incremental cost effectiveness of providing routine testing using one of the four interventions to all patients newly diagnosed with suspected alcohol-related liver fibrosis who might otherwise be referred for biopsy or specialist care on the basis of the clinical history and physical examination and/or standard liver function tests.

**5. Report methods for synthesis of evidence of clinical effectiveness**

Systematic reviews of the evidence for diagnostic accuracy and health and quality of life outcomes will be undertaken; these will be informed by the general principles recommended in the PRISMA (formerly QUOROM) statement.[10] Evidence of diagnostic accuracy will be sought from studies which compare any of the four interventions with detected pathology or other diagnostic tools. Sensitivity (the proportion of true positives) and specificity (the proportion of true negatives) will be assessed.

In addition to the formal systematic review, the manufacturers may provide unpublished and confidential data, which would be analysed to provide further information on test characteristics.

The description of studies below covers studies that would provide direct comparative evidence for outcomes of interest. The Assessment Team recognizes that such studies are unlikely to exist and that indirect evidence will be needed to fill in the data requirements of the model. These data will be sought as the model design becomes apparent using the appropriate criteria. The same sources will apply.

*5.1 Population*

- Inclusion criteria: Patients with suspected liver fibrosis related to alcohol consumption.
- Exclusion criteria: Liver dysfunction attributed to other possible aetiologies. However, if time and evidence permit, consideration will be given to patients with suspected liver fibrosis related to alcohol consumption who also have hepatitis C.

*5.2 Interventions*

- Enhanced Liver Fibrosis (ELF) blood test
- FibroTest blood test
- FibroMax blood test
- FibroScan (transient elastography)

*5.3 Comparators*

Referral to specialty care or biopsy based on clinical suspicion of liver fibrosis based on symptoms and/or liver function test results.

*5.4 Outcomes*

- Diagnostic test accuracy
- Number of patients requiring referral to secondary care
- Number of patients requiring liver biopsy
- Number of patients giving up alcohol, or significantly reducing alcohol consumption
- Long-term patient outcomes (disease progression, complications related to liver disease, need for liver transplantation, mortality)
- Adverse effects of testing
- Health-related quality of life

*5.5 Study design*

- Inclusion criteria: for the review of clinical effectiveness the best available level of evidence will be included, with priority given to controlled studies if available. However, this criterion will be relaxed for the consideration of adverse events, for which observational studies may be included even if controlled studies are available.
- Exclusion criteria: studies will be excluded if they do not meet the inclusion criteria, appear to be methodologically unsound, or do not report results in the necessary detail. The following will also be excluded:
    - Animal models
    - Preclinical and biological studies
    - Narrative reviews, editorials and opinions
    - Reports published as meeting abstracts only, where insufficient methodological details are reported to allow critical appraisal of study quality.

*5.6 Search strategy*

The search strategy will comprise the following main elements:

- Searching of electronic databases
- Contact with experts in the field
- Scrutiny of bibliographies of retrieved papers.

The electronic databases to be searched will include MEDLINE; Medline in Process; EMBASE; the Cochrane Database of Systematic Reviews, and the Cochrane Controlled Trials Register. A draft Medline search strategy is included in Appendix 1.

All citations will be imported into Reference Manager software and screened for inclusion on the basis of the inclusion/exclusion criteria listed above. Screening will be done in three stages,

sifting first by title, then by abstract, and finally by full text, excluding at each step studies which do not satisfy the inclusion/exclusion criteria.

### *5.7 Data extraction strategy*

Data will be extracted by one researcher using a standardised data extraction form. Any studies which give rise to uncertainty will be reviewed by a second researcher, and any disagreements will be resolved by discussion, with involvement of a third researcher where necessary.

### *5.8 Quality assessment strategy*

The nature of the quality assessment which will be undertaken will depend on the types of studies identified, but will be undertaken using appropriate and established tools (eg the QUADAS checklist for studies of diagnostic accuracy[11]).

### *5.9 Methods of analysis/synthesis*

Data will be tabulated and discussed in a narrative review. Where appropriate, meta-analysis will be employed to provide pooled estimates of test accuracy, and of patient outcomes.

### *5.10 Methods for estimating quality of life*

In order to reflect the chronic nature of the disease, the time horizon of the analysis will be a patient's lifetime. The perspective will be that of the National Health Services and Personal Social Services. Both cost and QALY will be discounted at 3.5% as recommended by NICE.

## 6. Report methods for synthesising evidence of cost effectiveness

A systematic review of the existing literature studying the cost effectiveness of non-invasive diagnostic assessment tools for the detection of liver fibrosis will be undertaken.

### *6.1 Identifying and systematically reviewing published cost effectiveness studies*

Studies relating to cost effectiveness will be identified using an economic search filter which will be integrated into the search strategy detailed in Section 5.6. This economic search filter is presented in Appendix 1.

### *6.2 Evaluation of costs and cost effectiveness*

The quality of identified economic literature will be assessed using a combination of key components of the British Medical Journal checklist for economic evaluations[12] together with the Eddy checklist on mathematical models[13] (see Appendix 2).

### *6.3 Development of a health economic model*

A *de novo* economic evaluation of the cost effectiveness of the use of each of the four interventions will be conducted. A model will be developed to identify whether the routine testing of all patients with one (or if resources allow multiple) non-invasive diagnostic test(s)

who are suspected of having alcohol-related liver disease and who would be referred for a liver biopsy is a cost effective use of resources.

The primary outcome from the model will be an estimate of the incremental cost per additional quality-adjusted life year (QALY) gained associated with the use of non-invasive diagnostic tests in the assessment of alcohol-related liver disease. A lifetime time horizon will be used in order to reflect the chronic effects of alcohol-related liver disease and potential mortality. The perspective used will be that of the UK National Health Service and Personal Social Services. Costs and QALYs will be discounted at 3.5% as recommended in the NICE reference case.[14] Modelling assumptions will be taken from the literature, supplemented by clinical expert opinion where required.

The development of the model is likely to be an iterative process. A conceptual model will be developed in conjunction with clinical experts to capture the current pathway of care for patients with suspected alcohol liver disease, and furthermore, how this pathway would change should non-invasive diagnostic tests become available for routine use. The conceptual model will indicate the data requirements which will be sought both from the published literature and within commercial in confidence data held by the manufacturers. The model is likely to evolve following discussions with project stakeholders and the Diagnostics Advisory Committee, and according to the availability of data.

Ideally, health related quality of life evidence will be available directly from the review literature. In the absence of such evidence, the mathematical model may use indirect evidence on quality of life from alternative sources. Quality of life data will be reviewed and used to generate the quality adjustment weights required for the model. In addition to the reviewed literature, national sources (e.g. NHS reference costs, national unit costs,[15] British National Formulary) will be used to estimate resource use and costs for use in the economic model.

It is anticipated that there may be limited evidence for some of the parameters that will be included in the economic model. Therefore, the uncertainty around the parameter estimates will be modelled to take this into account. The uncertainty in the input parameters will be propagated through the model using PSA to characterise uncertainty in the outputs. Results will include the presentation of a cost effectiveness acceptability curve and the reporting of the expected value of perfect information.[16] If resources allow, the cost effectiveness of collecting further information will be explicitly explored using Expected Value of Partial Perfect Information[17] or the Expected Value of Sample Information techniques[18] which the team have experience of undertaking.[19,20]

## 7. Handling information from the companies

All data submitted by the manufacturers/sponsors will be considered if received by the Assessment Team in a timely manner. Data arriving after this date will not be considered. Data which meet the inclusion criteria for the review will be extracted and quality assessed in accordance with the procedures outlined in this protocol.

Any 'commercial in confidence' data taken from a company submission will be underlined in the assessment report (followed by an indication of the relevant company name e.g. in brackets) presented to the Diagnostics Advisory Committee only. In the version of the report released to manufacturers and other stakeholders, commercial in confidence data will be blacked out, thus ensuring confidentiality.

## 8. Competing interests of authors

None.

## 9. Timetable/milestones

The dates in this section are dependent on NICE's agreement to hold three Committee meetings (in May, September and November 2010) to discuss the pilot topic.

| Milestone | Date to be completed |
|---|---|
| Draft protocol | 11 December 2009 |
| Final protocol | 22 December 2009 |
| Progress report | Weekly meetings |
| Draft assessment report to NICE for Committee consideration | 7th May 2010 |
| Presentation of draft assessment report, including model, to Diagnostics Advisory Committee (1st meeting) | 28th May 2010 |
| Final assessment report to NICE for circulation to stakeholders | 10 weeks before 2nd Committee meeting (i.e. early – mid July 2010) |
| Stakeholder comments to NICE | July – August 2010 |
| 2nd Diagnostics Advisory Committee meeting | Late September 2010 |
| 3rd Diagnostics Advisory Committee meeting | Late November 2010 (8 weeks after 2nd Committee meeting) |

## 10. Appendices

### Appendix 1: Draft Medline search strategy and economic search filter

### OVID Medline or Medline in Process

1. (enhanced adj liver adj fibrosis).tw.
2. (elf adj test$).tw.
3. (elf and diagnos$).tw.
4. (elf and (fibros*s or cirrhos*s).tw.
5. elf.tw.
6. exp liver cirrhosis/ or exp liver diseases, alcoholic/
7. 5 and 6
8. 1 or 2 or 3 or 4 or 7
9. fibrotest.tw.
10. fibrosure.tw.
11. fibromax.tw.
12. ashtest.tw.
13. fibroscan.tw.
14. (transient adj elastograph$).tw.
15. (elastograph$ and liver).tw.
16. or/9 to 15
17. exp liver cirrhosis/ or exp liver diseases, alcoholic/
18. (fibros*s or cirrhos*s).tw.
19. 17 or 18
20. Biological Markers/
21. biomarker$.tw.
22. (marker$ and (biologic$ or biochemical or serum or direct or indirect)).tw.
23. Algorithms/
24. algorithm$.tw.
25. (composite and blood).tw.
26. or/20-25
27. 19 and 26
28. Hyaluronic Acid/
29. ((hyaluronic adj acid) or (hyalauronate or hyaluronan).tw.
30. 28 or 29
31. ((procollagen or piiinp or p3np or ppcp)).tw.
32. ((tissue and inhibitor and metalloproteinase$) or timps).tw.
33. 30 and 31 and 32
34. 30 or 31 or 32
35. 34 and 19
36. Alpha-Macroglobulins/
37. ((alpha and macroglobulin$) or (alpha adj 2m)).tw.
38. or/36-37
39. ((apolipoprotein$ adj a1) or apoa1).tw.
40. Haptoglobins/
41. haptoglobin$.tw.
42. 40 or 41
43. (bilirubin$ or hematoidin$).tw.

44. (gamma adj glutamyl adj transpeptidase$).tw.
45. (gamma adj glutamyltransferase$).tw.
46. ((gamma adj gt) or ggt or ggtp).tw.
47. or/44-46
48. 38 and 39 and 42 and 43 and 47
49. 38 or 39 or 42 or 43 or 47
50. 49 and 19
51. ((alanine adj aminotransferase$) or aminotransaminase$).tw.
52. (serum adj glutamic adj pyruvic adj transaminase).tw.
53. sgpt.tw
54. or/51-53
55. (aspartate adj (aminotransferase$ or aminotransaminase$)).tw.
56. (serum adj glutamic adj oxaloacetic adj transaminase$).tw.
57. sgot.tw
58. or/55-57
59. 38 and 39 and 42 and 43 and 47 and 54 and 58
60. 38 or 39 or 42 or 43 or 47 or 54 or 58
61. 60 and 19
62. exp "Sensitivity and Specificity"/
63. sensitivity.tw.
64. specificity.tw.
65. ((pre-test or pretest) adj probability).tw.
66. post-test probability.tw.
67. predictive value$.tw.
68. likelihood ratio$.tw.
69. or/62-68
70. 27 and 69
71. 35 and 69
72. 50 and 69
73. 61 and 69
74. or/70-73
75. iqur.tw.
76. biopredictive.tw.
77. echosens.tw.
78. or/75-77
79. 8 or 16 or 33 or 48 or 74 or 78

**Econometric search filter (OVID Medline) to follow from the above searches**

1. exp "costs and cost analysis"/
2. economics/
3. exp economics, hospital/
4. exp economics, medical/
5. economics, nursing/
6. exp models, economic/
7. economics, pharmaceutical/
8. exp "fees and charges"/
9. exp budgets/
10. budget$.tw
11. ec.fs

12. cost$.ti
13. (cost$ adj2 (effective$ or utilit$ or benefit$ or minimi$)).ab
14. (economic$ or pharmacoeconomic$ or pharmaco-economic$).ti
15. (price$ or pricing$).tw
16. (financial or finance or finances or financed).tw
17. (fee or fees).tw
18. (value adj2 (money or monetary)).tw
19. quality-adjusted life years/
20. (qaly or qalys).af.
21. (quality adjusted life year or quality adjusted life years).af.
22. or/1-21
23. 22 and 79 (above).

## Appendix 2: Critical appraisal checklist for economic evaluations using key components of the British Medical Journal checklist for economic evaluations[21] together with the Eddy checklist on mathematical models employed in technology assessments[13]

| | | |
|---|---|---|
| Reference ID | | |
| Title | | |
| Authors | | |
| Year | | |
| **Modelling assessments should include:** | | **Yes/No** |
| 1 | A statement of the problem; | |
| 2 | A discussion of the need for modelling vs. alternative methodologies | |
| 3 | A description of the relevant factors and outcomes; | |
| 4 | A description of the model including reasons for this type of model and a specification of the scope including; time frame, perspective, comparators and setting. *Note: n=number of health states within sub-model* | |
| 5 | A description of data sources (including subjective estimates), with a description of the strengths and weaknesses of each source, with reference to a specific classification or hierarchy of evidence; | |
| 6 | A list of assumptions pertaining to: the structure of the model (e.g. factors included, relationships, and distributions) and the data; | |
| 7 | A list of parameter values that will be used for a base case analysis, and a list of the ranges in those values that represent appropriate confidence limits and that will be used in a sensitivity analysis; | |
| 8 | The results derived from applying the model for the base case; | |
| 9 | The results of the sensitivity analyses; unidimensional; best/worst case; multidimensional (Monte Carlo/parametric); threshold. | |
| 10 | A discussion of how the modelling assumptions might affect the results, indicating both the direction of the bias and the approximate magnitude of the effect; | |
| 11 | A description of the validation undertaken including; concurrence of experts; internal consistency; external consistency; predictive validity. | |
| 12 | A description of the settings to which the results of the analysis can be applied and a list of factors that could limit the applicability of the results; | |
| 13 | A description of research in progress that could yield new data that could alter the results of the analysis | |

Reference List


1. Mann, R. E., Smart, R. G., and Govoni, R. The epidemiology of alcoholic liver disease. *Alcohol Research & Health* 2003; **27** 209-219.

2. Marsano, L. S., Mendez, C., Hill, D., Barve, S., and McClain, C. J. Diagnosis and treatment of alcoholic liver disease and its complications. *Alcohol Research & Health* 2003; **27** 247-256.

3. Walsh, K. and Alexander, G. Alcoholic liver disease. *Postgraduate Medical Journal* 2000; **76** 280-286.

4. Williams, R. The pervading influence of alcoholic liver disease in hepatology. *Alcohol & Alcoholism* 2008; **43** 393-397.

5. The National Clinical Guideline Centre for acute and chronic conditions Alcohol use disorders: diagnosis and clinical management of alcohol-related physical complications. NICE guideline. Draft for consultation. 2009;

6. Fleming, K. M., Aithal, G. P., Solaymani-Dodaran, M., Card, T. R., and West, J. Incidence and prevalence of cirrhosis in the United Kingdom, 1992-2001: a general population-based study. *Journal of Hepatology* 2008; **49** 732-738.

7. British Liver Trust Fighting liver disease. *Internet* 2009;

8. Fuller, E., Jotangia, D., and Farrell, M. Alcohol misuse and dependence. 2009;

9. Statistics for Wales Welsh Health Survey 2008. 2009;

10. Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. G. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009; **339** 332-339.

11. Whiting, P., Rutjes, A. W., Reitsma, J. B., Bossuyt, P. M., and Kleijnen, J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology* 2003; **3** 25-

12. Drummond, M. and Jefferson, T. O. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;275-283.

13. Eddy, D. M. The role of mathematical modeling. 1985;144-154.

14. National Institute for Health and Clinical Excellence Guide to the methods of technology appraisals. *Internet* 2008;

15. Curtis, L. Unit costs of health and social care. 2008;

16. Claxton, K. and Posnett, J. An economic approach to clinical trial design and research priority-setting. *Health Economics* 1996; **5** 513-524.

17. Felli, J. C. and Hazen, G. B. Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* 1998; **18** 95-109.

18. Ades, A. E., Lu, G., and Claxton, K. Expected values of sample information calculation in medical decision making. *Medical Decision Making* 2004; **24** 207-227.

19. Stevenson, M. D., Oakley, J. E., Lloyd Jones, M., Brennan, A., Compston, J. E., McCloskey, E. V., and Selby, P. L. The cost-effectiveness of an RCT to establish whether 5 or 10 years of bisphosphonate treatment is the better duration for women with a prior fracture. *Medical Decision Making* 2009; **29** 678-689.

20. Stevenson, M. D. and Lloyd Jones, M. Vitamin K to prevent fractures in older women: a systematic review and economic evaluation. *Health Technology Assessment* 2009; **13(45)**

21. Drummond, M and Jefferson, TO Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996; **313** 275-283.

# Health Technology Assessment programme

## HTA Commissioning Board *(continued)*

Professor Barney Reeves,
Professorial Research Fellow
in Health Services Research,
Department of Clinical Science,
University of Bristol

Professor Martin Underwood,
Professor of Primary Care
Research, Warwick Medical
School, University of Warwick

Professor Marion Walker,
Professor in Stroke Rehabilitation,
Associate Director UK Stroke
Research Network, University of
Nottingham

Dr Duncan Young,
Senior Clinical Lecturer and
Consultant, Nuffield Department
of Anaesthetics, University of
Oxford

### *Observers*

Dr Tom Foulks,
Medical Research Council

Dr Kay Pattison,
Senior NIHR Programme
Manager, Department of Health

## HTA Clinical Evaluation and Trials Board

**Chair,**
**Professor Sallie Lamb,**
Director,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick and Professor of
Rehabilitation,
Nuffield Department of Orthopaedic,
Rheumatology and Musculoskeletal Sciences,
University of Oxford

**Deputy Chair,**
**Professor Jenny Hewison,**
Professor of the Psychology of Health Care,
Leeds Institute of Health Sciences,
University of Leeds

**Programme Director,**
**Professor Tom Walley, CBE,**
Director, NIHR HTA programme, Professor of
Clinical Pharmacology, University of Liverpool

### *Members*

Professor Keith Abrams,
Professor of Medical Statistics,
Department of Health Sciences,
University of Leicester

Professor Martin Bland,
Professor of Health Statistics,
Department of Health Sciences,
University of York

Professor Jane Blazeby,
Professor of Surgery and
Consultant Upper GI Surgeon,
Department of Social Medicine,
University of Bristol

Professor Julia M Brown,
Director, Clinical Trials Research
Unit, University of Leeds

Professor Alistair Burns,
Professor of Old Age Psychiatry,
Psychiatry Research Group, School
of Community-Based Medicine,
The University of Manchester &
National Clinical Director for
Dementia, Department of Health

Dr Jennifer Burr,
Director, Centre for Healthcare
Randomised trials (CHART),
University of Aberdeen

Professor Linda Davies,
Professor of Health Economics,
Health Sciences Research Group,
University of Manchester

Professor Simon Gilbody,
Prof of Psych Medicine and Health
Services Research, Department of
Health Sciences, University of York

Professor Steven Goodacre,
Professor and Consultant in
Emergency Medicine, School of
Health and Related Research,
University of Sheffield

Professor Dyfrig Hughes,
Professor of Pharmacoeconomics,
Centre for Economics and Policy
in Health, Institute of Medical
and Social Care Research, Bangor
University

Professor Paul Jones,
Professor of Respiratory Medicine,
Department of Cardiac and
Vascular Science, St George's
Hospital Medical School,
University of London

Professor Khalid Khan,
Professor of Women's Health and
Clinical Epidemiology, Barts and
the London School of Medicine,
Queen Mary, University of London

Professor Richard J McManus,
Professor of Primary Care
Cardiovascular Research, Primary
Care Clinical Sciences Building,
University of Birmingham

Professor Helen Rodgers,
Professor of Stroke Care, Institute
for Ageing and Health, Newcastle
University

Professor Ken Stein,
Professor of Public Health,
Peninsula Technology Assessment
Group, Peninsula College
of Medicine and Dentistry,
Universities of Exeter and
Plymouth

Professor Jonathan Sterne,
Professor of Medical Statistics
and Epidemiology, Department
of Social Medicine, University of
Bristol

Mr Andy Vail,
Senior Lecturer, Health Sciences
Research Group, University of
Manchester

Professor Clare Wilkinson,
Professor of General Practice and
Director of Research North Wales
Clinical School, Department of
Primary Care and Public Health,
Cardiff University

Dr Ian B Wilkinson,
Senior Lecturer and Honorary
Consultant, Clinical Pharmacology
Unit, Department of Medicine,
University of Cambridge

### *Observers*

Ms Kate Law,
Director of Clinical Trials,
Cancer Research UK

Dr Morven Roberts,
Clinical Trials Manager, Health
Services and Public Health
Services Board, Medical Research
Council

## Diagnostic Technologies and Screening Panel

### Members

**Chair,**
**Professor Lindsay Wilson Turnbull,**
Scientific Director of the Centre for Magnetic Resonance Investigations and YCR Professor of Radiology, Hull Royal Infirmary

Professor Judith E Adams,
Consultant Radiologist, Manchester Royal Infirmary, Central Manchester & Manchester Children's University Hospitals NHS Trust, and Professor of Diagnostic Radiology, University of Manchester

Mr Angus S Arunkalaivanan,
Honorary Senior Lecturer, University of Birmingham and Consultant Urogynaecologist and Obstetrician, City Hospital, Birmingham

Dr Diana Baralle,
Consultant and Senior Lecturer in Clinical Genetics, University of Southampton

Dr Stephanie Dancer,
Consultant Microbiologist, Hairmyres Hospital, East Kilbride

Dr Diane Eccles,
Professor of Cancer Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital

Dr Trevor Friedman,
Consultant Liason Psychiatrist, Brandon Unit, Leicester General Hospital

Dr Ron Gray,
Consultant, National Perinatal Epidemiology Unit, Institute of Health Sciences, University of Oxford

Professor Paul D Griffiths,
Professor of Radiology, Academic Unit of Radiology, University of Sheffield

Mr Martin Hooper,
Public contributor

Professor Anthony Robert Kendrick,
Associate Dean for Clinical Research and Professor of Primary Medical Care, University of Southampton

Dr Nicola Lennard,
Senior Medical Officer, MHRA

Dr Anne Mackie,
Director of Programmes, UK National Screening Committee, London

Mr David Mathew,
Public contributor

Dr Michael Millar,
Consultant Senior Lecturer in Microbiology, Department of Pathology & Microbiology, Barts and The London NHS Trust, Royal London Hospital

Mrs Una Rennard,
Public contributor

Dr Stuart Smellie,
Consultant in Clinical Pathology, Bishop Auckland General Hospital

Ms Jane Smith,
Consultant Ultrasound Practitioner, Leeds Teaching Hospital NHS Trust, Leeds

Dr Allison Streetly,
Programme Director, NHS Sickle Cell and Thalassaemia Screening Programme, King's College School of Medicine

Dr Matthew Thompson,
Senior Clinical Scientist and GP, Department of Primary Health Care, University of Oxford

Dr Alan J Williams,
Consultant Physician, General and Respiratory Medicine, The Royal Bournemouth Hospital

### Observers

Dr Tim Elliott,
Team Leader, Cancer Screening, Department of Health

Dr Joanna Jenkinson,
Board Secretary, Neurosciences and Mental Health Board (NMHB), Medical Research Council

Professor Julietta Patnick,
Director, NHS Cancer Screening Programme, Sheffield

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

## Disease Prevention Panel

### Members

**Chair,**
**Professor Margaret Thorogood,**
Professor of Epidemiology, University of Warwick Medical School, Coventry

Dr Robert Cook,
Clinical Programmes Director, Bazian Ltd, London

Dr Colin Greaves,
Senior Research Fellow, Peninsula Medical School (Primary Care)

Mr Michael Head,
Public contributor

Professor Cathy Jackson,
Professor of Primary Care Medicine, Bute Medical School, University of St Andrews

Dr Russell Jago,
Senior Lecturer in Exercise, Nutrition and Health, Centre for Sport, Exercise and Health, University of Bristol

Dr Julie Mytton,
Consultant in Child Public Health, NHS Bristol

Professor Irwin Nazareth,
Professor of Primary Care and Director, Department of Primary Care and Population Sciences, University College London

Dr Richard Richards,
Assistant Director of Public Health, Derbyshire County Primary Care Trust

Professor Ian Roberts,
Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine

Dr Kenneth Robertson,
Consultant Paediatrician, Royal Hospital for Sick Children, Glasgow

Dr Catherine Swann,
Associate Director, Centre for Public Health Excellence, NICE

Mrs Jean Thurston,
Public contributor

Professor David Weller,
Head, School of Clinical Science and Community Health, University of Edinburgh

### Observers

Ms Christine McGuire,
Research & Development, Department of Health

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

## External Devices and Physical Therapies Panel

### Members

**Chair,**
**Dr John Pounsford,**
Consultant Physician North Bristol NHS Trust

**Deputy Chair,**
**Professor E Andrea Nelson,**
Reader in Wound Healing and Director of Research, University of Leeds

Professor Bipin Bhakta,
Charterhouse Professor in Rehabilitation Medicine, University of Leeds

Mrs Penny Calder,
Public contributor

Dr Dawn Carnes,
Senior Research Fellow, Barts and the London School of Medicine and Dentistry

Dr Emma Clark,
Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol

Mrs Anthea De Barton-Watson,
Public contributor

Professor Nadine Foster,
Professor of Musculoskeletal Health in Primary Care Arthritis Research, Keele University

Dr Shaheen Hamdy,
Clinical Senior Lecturer and Consultant Physician, University of Manchester

Professor Christine Norton,
Professor of Clinical Nursing Innovation, Bucks New University and Imperial College Healthcare NHS Trust

Dr Lorraine Pinnigton,
Associate Professor in Rehabilitation, University of Nottingham

Dr Kate Radford,
Senior Lecturer (Research), University of Central Lancashire

Mr Jim Reece,
Public contributor

Professor Maria Stokes,
Professor of Neuromusculoskeletal Rehabilitation, University of Southampton

Dr Pippa Tyrrell,
Senior Lecturer/Consultant, Salford Royal Foundation Hospitals' Trust and University of Manchester

Dr Nefyn Williams,
Clinical Senior Lecturer, Cardiff University

### Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

## Interventional Procedures Panel

### Members

**Chair,**
**Professor Jonathan Michaels,**
Professor of Vascular Surgery, University of Sheffield

**Deputy Chair,**
**Mr Michael Thomas,**
Consultant Colorectal Surgeon, Bristol Royal Infirmary

Mrs Isabel Boyer,
Public contributor

Mr Sankaran Chandra Sekharan,
Consultant Surgeon, Breast Surgery, Colchester Hospital University NHS Foundation Trust

Professor Nicholas Clarke,
Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust

Ms Leonie Cooke,
Public contributor

Mr Seumas Eckford,
Consultant in Obstetrics & Gynaecology, North Devon District Hospital

Professor Sam Eljamel,
Consultant Neurosurgeon, Ninewells Hospital and Medical School, Dundee

Dr Adele Fielding,
Senior Lecturer and Honorary Consultant in Haematology, University College London Medical School

Dr Matthew Hatton,
Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust

Dr John Holden,
General Practitioner, Garswood Surgery, Wigan

Dr Fiona Lecky,
Senior Lecturer/Honorary Consultant in Emergency Medicine, University of Manchester/Salford Royal Hospitals NHS Foundation Trust

Dr Nadim Malik,
Consultant Cardiologist/Honorary Lecturer, University of Manchester

Mr Hisham Mehanna,
Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust

Dr Jane Montgomery,
Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust

Professor Jon Moss,
Consultant Interventional Radiologist, North Glasgow Hospitals University NHS Trust

Dr Simon Padley,
Consultant Radiologist, Chelsea & Westminster Hospital

Dr Ashish Paul,
Medical Director, Bedfordshire PCT

Dr Sarah Purdy,
Consultant Senior Lecturer, University of Bristol

Dr Matthew Wilson,
Consultant Anaesthetist, Sheffield Teaching Hospitals NHS Foundation Trust

Professor Yit Chiun Yang,
Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust

### Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

## Pharmaceuticals Panel

### Members

**Chair,**
**Professor Imti Choonara,**
Professor in Child Health, University of Nottingham

**Deputy Chair,**
**Dr Yoon K Loke,**
Senior Lecturer in Clinical Pharmacology, University of East Anglia

Dr Martin Ashton-Key,
Medical Advisor, National Commissioning Group, NHS London

Dr Peter Elton,
Director of Public Health, Bury Primary Care Trust

Dr Ben Goldacre,
Research Fellow, Epidemiology London School of Hygiene and Tropical Medicine

Dr James Gray,
Consultant Microbiologist, Department of Microbiology, Birmingham Children's Hospital NHS Foundation Trust

Dr Jurjees Hasan,
Consultant in Medical Oncology, The Christie, Manchester

Dr Carl Heneghan,
Deputy Director Centre for Evidence-Based Medicine and Clinical Lecturer, Department of Primary Health Care, University of Oxford

Dr Dyfrig Hughes,
Reader in Pharmacoeconomics and Deputy Director, Centre for Economics and Policy in Health, IMSCaR, Bangor University

Dr Maria Kouimtzi,
Pharmacy and Informatics Director, Global Clinical Solutions, Wiley-Blackwell

Professor Femi Oyebode,
Consultant Psychiatrist and Head of Department, University of Birmingham

Dr Andrew Prentice,
Senior Lecturer and Consultant Obstetrician and Gynaecologist, The Rosie Hospital, University of Cambridge

Ms Amanda Roberts,
Public contributor

Dr Gillian Shepherd,
Director, Health and Clinical Excellence, Merck Serono Ltd

Mrs Katrina Simister,
Assistant Director New Medicines, National Prescribing Centre, Liverpool

Professor Donald Singer,
Professor of Clinical Pharmacology and Therapeutics, Clinical Sciences Research Institute, CSB, University of Warwick Medical School

Mr David Symes,
Public contributor

Dr Arnold Zermansky,
General Practitioner, Senior Research Fellow, Pharmacy Practice and Medicines Management Group, Leeds University

### Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Mr Simon Reeve,
Head of Clinical and Cost-Effectiveness, Medicines, Pharmacy and Industry Group, Department of Health

Dr Heike Weber,
Programme Manager, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

## Psychological and Community Therapies Panel

### Members

**Chair,**
**Professor Scott Weich,**
Professor of Psychiatry, University of Warwick, Coventry

**Deputy Chair,**
**Dr Howard Ring,**
Consultant & University Lecturer in Psychiatry, University of Cambridge

Professor Jane Barlow,
Professor of Public Health in the Early Years, Health Sciences Research Institute, Warwick Medical School

Dr Sabyasachi Bhaumik,
Consultant Psychiatrist, Leicestershire Partnership NHS Trust

Mrs Val Carlill,
Public contributor

Dr Steve Cunningham,
Consultant Respiratory Paediatrician, Lothian Health Board

Dr Anne Hesketh,
Senior Clinical Lecturer in Speech and Language Therapy, University of Manchester

Dr Peter Langdon,
Senior Clinical Lecturer, School of Medicine, Health Policy and Practice, University of East Anglia

Dr Yann Lefeuvre,
GP Partner, Burrage Road Surgery, London

Dr Jeremy J Murphy,
Consultant Physician and Cardiologist, County Durham and Darlington Foundation Trust

Dr Richard Neal,
Clinical Senior Lecturer in General Practice, Cardiff University

Mr John Needham,
Public contributor

Ms Mary Nettle,
Mental Health User Consultant

Professor John Potter,
Professor of Ageing and Stroke Medicine, University of East Anglia

Dr Greta Rait,
Senior Clinical Lecturer and General Practitioner, University College London

Dr Paul Ramchandani,
Senior Research Fellow/Cons. Child Psychiatrist, University of Oxford

Dr Karen Roberts,
Nurse/Consultant, Dunston Hill Hospital, Tyne and Wear

Dr Karim Saad,
Consultant in Old Age Psychiatry, Coventry and Warwickshire Partnership Trust

Dr Lesley Stockton,
Lecturer, School of Health Sciences, University of Liverpool

Dr Simon Wright,
GP Partner, Walkden Medical Centre, Manchester

### Observers

Dr Kay Pattison,
Senior NIHR Programme Manager, Department of Health

Dr Morven Roberts,
Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council

Professor Tom Walley, CBE,
Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool

Dr Ursula Wells,
Principal Research Officer, Policy Research Programme, Department of Health

# Expert Advisory Network

## *Members*

Professor Douglas Altman,
Professor of Statistics in Medicine,
Centre for Statistics in Medicine,
University of Oxford

Professor John Bond,
Professor of Social Gerontology
& Health Services Research,
University of Newcastle upon Tyne

Professor Andrew Bradbury,
Professor of Vascular Surgery,
Solihull Hospital, Birmingham

Mr Shaun Brogan,
Chief Executive, Ridgeway
Primary Care Group, Aylesbury

Mrs Stella Burnside OBE,
Chief Executive, Regulation and
Improvement Authority, Belfast

Ms Tracy Bury,
Project Manager, World
Confederation of Physical Therapy,
London

Professor Iain T Cameron,
Professor of Obstetrics and
Gynaecology and Head of the
School of Medicine, University of
Southampton

Professor Bruce Campbell,
Consultant Vascular & General
Surgeon, Royal Devon & Exeter
Hospital, Wonford

Dr Christine Clark,
Medical Writer and Consultant
Pharmacist, Rossendale

Professor Collette Clifford,
Professor of Nursing and Head
of Research, The Medical School,
University of Birmingham

Professor Barry Cookson,
Director, Laboratory of Hospital
Infection, Public Health
Laboratory Service, London

Dr Carl Counsell,
Clinical Senior Lecturer in
Neurology, University of Aberdeen

Professor Howard Cuckle,
Professor of Reproductive
Epidemiology, Department
of Paediatrics, Obstetrics &
Gynaecology, University of Leeds

Professor Carol Dezateux,
Professor of Paediatric
Epidemiology, Institute of Child
Health, London

Mr John Dunning,
Consultant Cardiothoracic
Surgeon, Papworth Hospital NHS
Trust, Cambridge

Mr Jonothan Earnshaw,
Consultant Vascular Surgeon,
Gloucestershire Royal Hospital,
Gloucester

Professor Martin Eccles,
Professor of Clinical Effectiveness,
Centre for Health Services
Research, University of Newcastle
upon Tyne

Professor Pam Enderby,
Dean of Faculty of Medicine,
Institute of General Practice
and Primary Care, University of
Sheffield

Professor Gene Feder,
Professor of Primary Care
Research & Development, Centre
for Health Sciences, Barts and The
London School of Medicine and
Dentistry

Mr Leonard R Fenwick,
Chief Executive, Freeman
Hospital, Newcastle upon Tyne

Mrs Gillian Fletcher,
Antenatal Teacher and Tutor and
President, National Childbirth
Trust, Henfield

Professor Jayne Franklyn,
Professor of Medicine, University
of Birmingham

Mr Tam Fry,
Honorary Chairman, Child
Growth Foundation, London

Professor Fiona Gilbert,
Consultant Radiologist and NCRN
Member, University of Aberdeen

Professor Paul Gregg,
Professor of Orthopaedic Surgical
Science, South Tees Hospital NHS
Trust

Bec Hanley,
Co-director, TwoCan Associates,
West Sussex

Dr Maryann L Hardy,
Senior Lecturer, University of
Bradford

Mrs Sharon Hart,
Healthcare Management
Consultant, Reading

Professor Robert E Hawkins,
CRC Professor and Director of
Medical Oncology, Christie CRC
Research Centre, Christie Hospital
NHS Trust, Manchester

Professor Richard Hobbs,
Head of Department of Primary
Care & General Practice,
University of Birmingham

Professor Alan Horwich,
Dean and Section Chairman,
The Institute of Cancer Research,
London

Professor Allen Hutchinson,
Director of Public Health and
Deputy Dean of ScHARR,
University of Sheffield

Professor Peter Jones,
Professor of Psychiatry, University
of Cambridge, Cambridge

Professor Stan Kaye,
Cancer Research UK Professor of
Medical Oncology, Royal Marsden
Hospital and Institute of Cancer
Research, Surrey

Dr Duncan Keeley,
General Practitioner (Dr Burch &
Ptnrs), The Health Centre, Thame

Dr Donna Lamping,
Research Degrees Programme
Director and Reader in
Psychology, Health Services
Research Unit, London School of
Hygiene and Tropical Medicine,
London

Professor James Lindesay,
Professor of Psychiatry for the
Elderly, University of Leicester

Professor Julian Little,
Professor of Human Genome
Epidemiology, University of
Ottawa

Professor Alistaire McGuire,
Professor of Health Economics,
London School of Economics

Professor Neill McIntosh,
Edward Clark Professor of Child
Life and Health, University of
Edinburgh

Professor Rajan Madhok,
Consultant in Public Health, South
Manchester Primary Care Trust

Professor Sir Alexander Markham,
Director, Molecular Medicine
Unit, St James's University
Hospital, Leeds

Dr Peter Moore,
Freelance Science Writer, Ashtead

Dr Andrew Mortimore,
Public Health Director,
Southampton City Primary Care
Trust

Dr Sue Moss,
Associate Director, Cancer
Screening Evaluation Unit,
Institute of Cancer Research,
Sutton

Professor Miranda Mugford,
Professor of Health Economics
and Group Co-ordinator,
University of East Anglia

Professor Jim Neilson,
Head of School of Reproductive
& Developmental Medicine
and Professor of Obstetrics
and Gynaecology, University of
Liverpool

Mrs Julietta Patnick,
Director, NHS Cancer Screening
Programmes, Sheffield

Professor Robert Peveler,
Professor of Liaison Psychiatry,
Royal South Hants Hospital,
Southampton

Professor Chris Price,
Director of Clinical Research,
Bayer Diagnostics Europe, Stoke
Poges

Professor William Rosenberg,
Professor of Hepatology and
Consultant Physician, University
of Southampton

Professor Peter Sandercock,
Professor of Medical Neurology,
Department of Clinical
Neurosciences, University of
Edinburgh

Dr Philip Shackley,
Senior Lecturer in Health
Economics, Sheffield Vascular
Institute, University of Sheffield

Dr Eamonn Sheridan,
Consultant in Clinical Genetics, St
James's University Hospital, Leeds

Dr Margaret Somerville,
Director of Public Health
Learning, Peninsula Medical
School, University of Plymouth

Professor Sarah Stewart-Brown,
Professor of Public Health,
Division of Health in the
Community, University of
Warwick, Coventry

Dr Nick Summerton,
GP Appraiser and Codirector,
Research Network, Yorkshire
Clinical Consultant, Primary Care
and Public Health, University of
Oxford

Professor Ala Szczepura,
Professor of Health Service
Research, Centre for Health
Services Studies, University of
Warwick, Coventry

Dr Ross Taylor,
Senior Lecturer, University of
Aberdeen

Dr Richard Tiner,
Medical Director, Medical
Department, Association of the
British Pharmaceutical Industry

Mrs Joan Webster,
Consumer Member, Southern
Derbyshire Community Health
Council

Professor Martin Whittle,
Clinical Co-director, National
Co-ordinating Centre for Women's
and Children's Health, Lymington

Current and past membership details of all HTA programme 'committees' are available from the HTA website (www.hta.ac.uk)

## Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

***We look forward to hearing from you.***

NETSCC, Health Technology Assessment
Alpha House
University of Southampton Science Park
Southampton SO16 7NS, UK
Email: hta@hta.ac.uk
www.hta.ac.uk

ISSN 1366-5278